FrameNet II: Extended Theory and Practice

Josef Ruppenhofer
Michael Ellsworth
Miriam R. L. Petruck
Christopher R. Johnson
Jan Scheffczyk

Printed September 14, 2010

Contents

Chapter 1

Introduction to the Project

The Berkeley FrameNet project is creating an on-line lexical resource for English, based on **frame semantics** and supported by corpus evidence. The aim is to document the range of semantic and syntactic combinatory possibilities– **valences**–of each word in each of its senses, through computer-assisted annotation of example sentences and automatic tabulation and display of the annotation results. The major product of this work, the FrameNet lexical database, currently contains more than 10,000 lexical units (defined below), more than 6,000 of which are fully annotated, in nearly 800 hierarchically-related semantic frames, exemplified in more than 135,000 annotated sentences. Beginning with Release 1.3, the quality of FrameNet data is monitored by a consistency management system. The database has gone through three releases, and is now in use by hundreds of researchers, teachers, and students around the world. (See the FrameNet Users page on our web-site). Active research projects are seeking to produce comparable frame-semantic lexicons for other languages and to devise means of automatically labeling running text with semantic frame information.

A **lexical unit (LU)** is a pairing of a word with a meaning. Typically, each sense of a polysemous word belongs to a different **semantic frame**, a script-like conceptual structure that describes a particular type of situation, object, or event along with its participants and props. For example, the Apply_heat frame describes a common situation involving a COOK, some FOOD, and a HEATING_INSTRUMENT, and is **evoked** by words such as *bake, blanch, boil, broil, brown, simmer, steam*, etc. We call these roles **frame elements (FEs)** and the frame-evoking words are LUs in the Apply_heat frame. Some frames are more abstract, such as Change_position_on_a_scale, which is evoked by LUs such as *decline, decrease, gain, plummet, rise*, etc., and has FEs such as ITEM, ATTRIBUTE, INITIAL_VALUE and FINAL_VALUE.

In the simplest case, the frame-evoking LU is a verb and the FEs are its syntactic dependents:

[$_{Cook}$ Matilde] **fried** [$_{Food}$ the catfish] [$_{Heating\_instrument}$ in a heavy iron skillet].

[$_{Item}$ Colgate's stock] **rose** [$_{Difference}$ $3.64] [$_{Final_value}$ to $49.94].

However, event nouns such as *reduction* in the Cause_change_of_scalar_position frame also evoke frames:

...the **reduction** [$_{Item}$ of debt levels] [$_{Value\_2}$ to $665 million] [$_{Value_1}$ from $2.6 billion]

or adjectives such as *asleep* in the Sleep frame:

[$_{Sleeper}$ They] [$_{Copula}$ were] **asleep** [$_{Duration}$ for hours]

The lexical entry for a predicating word, derived from such annotations, identifies the frame which underlies a given meaning and specifies the ways in which FEs are realized in structures headed by the word.

Many common nouns, such as artifacts like *hat* or *tower*, typically serve as **dependents** rather than clearly evoking their own frames. The main purpose of annotating such items is to identify the most common predicates that **govern** phrases headed by them, and thus to illustrate the ways in which these common nouns function as FEs within frames evoked by the governing predicates.

We **do** recognize that artifact and natural kind nouns also have a minimal frame structure of their own. For example, artifacts often occur together with expressions indicating their sub-type, the material of which they are made, their manner of production, and their purpose/use; these are defined as FEs in the frames for various types of artifacts. Consider two example sentences from the Clothing frame.

(1) He took a packet of Woodbines out of the breast pocket of [his WEARER] [cotton MATERIAL] [**shirt** GARMENT] and lit one.

(2) She had a [white DESCRIPTOR] [silk MATERIAL] [**blouse** GARMENT] on, and a severe grey skirt that reached halfway down her calves.

However, the frames evoked by artifact and natural kind nouns rarely dominate the clauses in which they occur, and so are seldom selected as targets of annotation.

Formally, FrameNet annotations are constellations of triples that make up the frame element realization for each annotated sentence, each consisting of a frame element (for example, **Food**), a **grammatical function** (say, **Object**) and a **phrase type** (say, **NP**). We think of these three types of annotation on each tagged frame element as *layers* and they are displayed as such in the annotation software used in the project. However, to avoid visual clutter, the grammatical function and phrase type layers are not displayed in the web-based report system. The full data, available as part of the data download (see the FNdata link on the FrameNet homepage), include these three layers (and several more not discussed here) for all of the annotated sentences, along with complete frame and FE descriptions, frame-to-frame relations, and lexical entries summarizing the valence patterns for each annotated LU.

FrameNet annotations derive from two sources. In pursuing the goal of recording the range of semantic and syntactic combinatory possibilities (valences) of each word in each of its senses, we normally concentrate on a particular target LU and extract sentences from the different texts of a corpus containing that LU. Then we annotate a selection of the extracted sentences in respect to the target LU. In another kind of work that represents a much smaller percentage of our overall annotations, we annotate running text. Full-text annotation differs from sentence annotation mostly in that the sentences are chosen for us, so to speak, by the author of the text. The annotation of running text is technically possible thanks to the annotation layering technique: FN lexicographers can one by one declare each word in a sentence a target, select a frame relative to which the new target is to be annotated, get a new set of annotation layers (frame element, grammatical function, phrase type) and appropriate frame element tags, and then annotate the relevant constituents.

1.1 Comparison with WordNet and ontologies

The FrameNet database is a lexical resource with unique characteristics that differentiate it from other resources such as commercially available dictionaries and thesauri as well as from the best-known on-line lexical resource, WordNet.

- Like dictionary subentries, FrameNet lexical units come with definitions, either from the Concise Oxford Dictionary, 10th Edition (courtesy of Oxford University Press) or a definition written by a FrameNet staff member.

- Unlike commercial dictionaries, we provide multiple annotated examples of each sense of a word (i.e. each lexical unit). Moreover, the set of examples (approximately 20 per LU) illustrates **all** of the combinatorial possibilities of the lexical unit.

- The examples are attestations taken from naturalistic corpora, rather than constructed by a linguist or lexicographer. The main FrameNet corpus is the 100-million-word British National CorpusBritish National Corpus (BNC), which is both large and balanced across genres (editorials, textbooks, advertisements, novels, sermons, etc.), but, of course, lacks many specifically American expressions. We also use U.S. newswire texts provided by the Linguistic Data ConsortiumLinguistic Data Consortium, and have recently acquired the newly released initial part of the American National Corpus, which we will begin using soon.

- Our analysis of the English lexicon proceeds frame by frame rather than by lemma, whereas traditional dictionary-making proceeds word by word through the alphabet. Thus, while a traditional lexicographer measures progress in words completed, FrameNet measures progress in frames completed. Having one or more LUs for a given word in completed frames does not preclude the possibility that there could be other LUs for the same word in future frames.

- Each lexical unit is linked to a semantic frame, and hence to the other words which evoke that frame. This makes the FrameNet database similar to a thesaurus, grouping together semantically similar words.

- WordNet and all ontologies provide some sort of hierarchical relations between their nodes; likewise, FrameNet includes a network of relations between frames. Several types are defined, of which the most important are:

  - Inheritance: An IS-A relation. The child frame is a subtype of the parent frame, and each FE in the parent is bound to a corresponding FE in the child. An example is the Revenge frame which inherits from the Rewards_and_punishments frame.
  - Using: The child frame presupposes the parent frame as background, e.g the Speed frame "uses" (or presupposes) the Motion frame; however, not all parent FEs need to be bound to child FEs.
  - Subframe: The child frame is a subevent of a complex event represented by the parent, e.g. the Criminal_process frame has subframes of Arrest, Arraignment, Trial, and Sentencing.
  - Perspective on: The child frame provides a particular perspective on an un-perspectivized parent frame. A pair of examples consists of the Hiring and Get_a_job frames, which perspectivize the Employment_start frame from the EMPLOYER's and the EMPLOYEE's point of view, respectively.

  These frame-to-frame relations are shown in the frame reports; the FE-to-FE relations are not shown in the frame reports but they can be viewed through the FrameGrapherFrameGrapher tool (accessible from the FrameNet homepage).

- Since we do not annotate many nouns denoting artifacts and natural kinds, the FrameNet database is not readily usable as an ontology of things. In this area, we mostly defer to WordNet, which provides extensive coverage, including hierarchical relations of areas such as animals, plants, etc.

1.2 What do we mean by *word*?

In this discussion, we have used the word *word* in talking about lexical units. The reality is actually rather complex. When we say that the word *bake* is polysemous, we mean that the lemma *bake.v* (which has the word-forms *bake, bakes, baked*, and *baking*) is linked to three different frames:

- Apply_heat: Michelle **baked** the potatoes for 45 minutes.
- Cooking_creation: Michelle **baked** her mother a cake for her birthday.
- Absorb_heat: The potatoes have to **bake** for more than 30 minutes.

These constitute three different LUs, with different definitions.

Multiword expressions such as *given name* and hyphenated words like *shut-eye* can also be LUs. Idiomatic phrases such as *middle of nowhere* and *give the slip (to)* are also defined as LUs in the appropriate frames (Isolated_places and Evading, respectively), and their internal structure is not analyzed.

For additional discussion also see the FAQs on the FrameNet website.

Chapter 2

Frame Development

The core of the process has always been looking at corpus attestations of a group of words that we believe to have some semantic overlap, and dividing these attestations into groups. Afterward, we combine the small groups into large enough groupings to make reasonable frames at which point we may (equivalently) call the words targets, lexical units, or frame-evoking elements. In the past, the criteria for such grouping have been informal and intuitive, but recently, the criteria have become more explicit. The smallest groups are formed as follows:

- All LUs in a frame must have the same **number** and **types** of frame elements in both explicit and implicit (NI) contexts.
  - Number

    If the number of essential, syntactically prominent frame elements differs from LU to LU or from sentence to sentence, this generally suggests that the frame should be split so that each resulting frame captures the difference. For instance, we have regularly split causatives from inchoatives. Consider the verbs *increase* and *diminish* in examples (1) to (4) below.

    (1) Also, the violent crime rate has **increased** from 455 to 563 offenses per 100,000 population, despite a decline in crime rates nationally.

    (2) The law has **increased** the Government Employee Bonus calculation rate from 6% to 7%.

    (3) From the summer 1998 to the summer 1999 the speed variation has **decreased**.

    (4) We have **decreased** the number of service calls.

    Examples (1) and (3) present the up or down changes in the value of a scalar attribute as events that happen by themselves, whereas examples (2) and (4) present them as caused by a Cause or an Agent. We have separated the two types of sentences into two different frames, an inchoative frame called Change_position_on_a_scale for examples (1) and (3) and a causative one called Cause_change_of_position_on_a_scale for examples (2) and (4).

    There may be a legitimate objection about the presence of an Agent or Cause being just a vague linguistic intuition and that we ignore the fact that everything that happens is caused. In this and in many other situations where one needs to decide between splitting and lumping causative and inchoative sentences, linguistic evidence suggests that splitting is the right way to proceed. First, there will typically be lexical units that exhibit only one of the two uses. For instance, the verb *gain* only has inchoative uses when referring to scalar change, while the verb *lower* only allows causative uses in the domain of scalar change. Second, cross-linguistic comparison also shows that other languages often distinguish inchoatives and causatives by derivational morphology.

    As pointed out above, we factor out language-wide constructions when figuring out how many participants there are in a particular use of a particular LU. Thus, we don't need an extra frame for the verb *sell* to account for uses like *Those boots sell well* because the sentence involves subject

selection construction (often called *middle*) which deprofiles and makes generic one or several of the prominent actors, in this case the SELLER. We then simply label SELLER as CNI , deeming it still conceptually present, but generic. There is indeed linguistic evidence for the conceptual presence of the SELLER: the notion of ease that is assessed requires a sentient being that could experience ease or difficulty selling.

Obviously, similar reasoning also applies to passives. In addition, we have used the idea that a construction can suppress the realization of a conceptually present frame element to account for the deleted objects in instructional imperatives, e.g. *Place __ in oven for twenty minutes.*

– Type

The basic semantic type for a frame element ought to be broadly constant across uses. If that is not so, it suggests the need to posit distinct frame elements. In some cases, however, we still want to recognize a relationship between frame elements whose syntactic form suggests that they refer to ontologically different kinds of entities. For example, in *I **want** [to win]* compared with *I **want** [an orange]*, both complements of the verb *want* have something to do with the Desiring frame, but each of the complements directly refers to something rather different. While the verb phrase *to win* refers to a type of event, the NP *an orange* refers to an entity. Clearly, the NP-complements of *want* and other such predicates like it, are understood to metonymically stand for events centrally involving them: 'to want an orange' is typically 'to want to eat an orange'. There is evidence that another event is to be understood by speaker and hearer when *want* is used with an object NP: sentences such as *I want your report on my desk tomorrow morning* include a time phrase that does not refer to the time of wanting but to the time when the desired event is to occur (or the desired state ought to hold).[1]

When, as with *want*, some of the target's complement types are appropriate to the ontological category of the frame element while others that are metonymically related are not, we normally create two distinct frame elements and relate them to each other via the Excludes relation (for which, see section 3.2.2.3).[2]

Splitting in this manner is useful in several ways. It allows the right frame element to frame element relations because in some cases, related frames will allow only one of the (metonymically or otherwise related) frame elements. Having distinct frame elements also facilitates finding data to study Pustejovskian coercions. And finally, it makes the annotation task easier if the clear ontological distinction is maintained as a frame element distinction.

- In aspectually complex frames, the lexical units should all entail the same set of stages and transitions. For instance, while *work on* and *develop* both evoke the background of a sentient AGENT expending effort to accomplish some GOAL, only the latter entails (in past tense utterances) that the GOAL was in fact accomplished. Similarly, the verb *shoot* can be used to report an event of firing a projectile at a person and hitting, but it does not entail that the person dies in contrast to, for instance, the verb *decapitate*. In all such cases, where lexical units differ in whether an intended and typical 'fulfillment' stage is in fact reached, we split frames. As with the number and type of frame elements, the requirement that the same subevents are entailed for all lexical units also argues for separating causatives and inchoatives, since the former class involves an extra causing subevent that the latter class lacks.

- The same frame elements will be profiled across all lexical units of a frame. That is, the same participant's point of view should be emphasized with all of them. Consider the difference between the verbs *buy* and *sell*: the first takes the BUYER's perspective and the second the SELLER's perspective.

  Note that when multiple perspectives on a type of scenario are possible, typically we relate the frames that carry particular perspectives to a non-perspectivized background/scenario frame via the Perspective_on relation (see Sect. 6.1.2). Thus, the perspectivized frames can be related to each other as candidates for paraphrasing via the background scenario. Figure 2.1 shows how the Commerce_buy and Commerce_sell frames are related to each other via the frame Commerce_goods-transfer, which is itself a sub-frame of the Commercial_transaction frame.

[1] There exists quite a bit of linguistic literature on the complementation of *want*. The same is true for many other issues that we will touch on. Within this manual, we will, however, not provide references in such cases since the amount of literature available is typically too vast to gather, let alone to summarize and evaluate.

[2] However, at this point, we cannot yet implement the concept of an abstract proto- or macro-frame element, which could generalize over the actually occurring FE types.

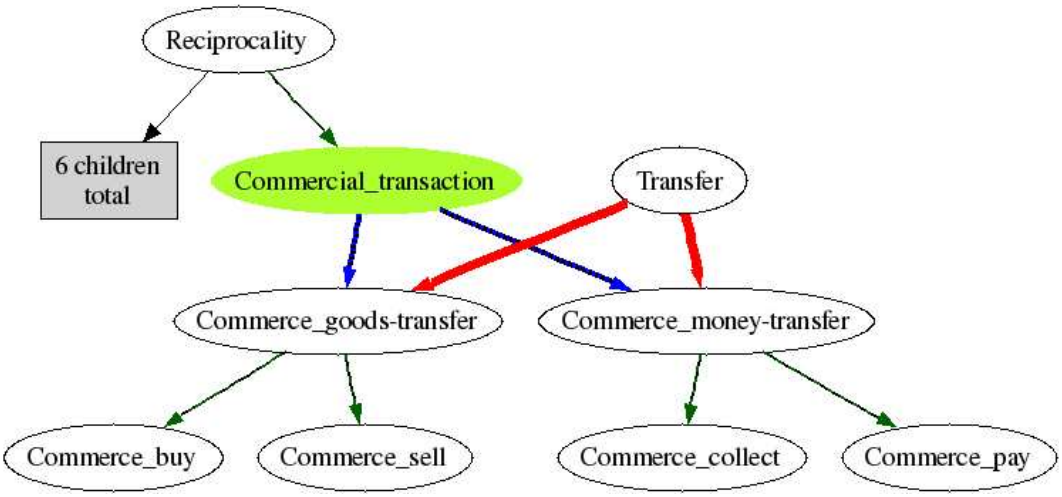

Figure 2.1: Commercial_Transaction

- We want the interrelations between frame elements to be the same for all the LUs in a frame. Consider, for instance, the fact that the PURPOSE frame elements of *buy* and *sell* have a relationship to different participants: a PURPOSE expressed with *buy* relates to the BUYER, a PURPOSE expressed with *sell* relates to the SELLER.

  (5) I **bought** this thing [in order to test it out].

  (6) At least you are **selling** them [in order to give the money to cancer research].

  The difference in the orientation of the PURPOSE frame element parallels the difference in perspective that we noted above and supports the need for a frame distinction for *buy* and *sell*.

- The presuppositions, expectations, and concomitants of the targets within a frame will be shared. For example, the verb *cross-examine* evokes a courtroom session, a preceding event of questioning by an opposing legal party, etc., differentiating it from the simpler *examine*. By similar reasoning, the verbs *receive* and *take* would be differentiated, since *receive* presupposes another willing agent participating as the DONOR and *take* does not.

- The basic denotation of the targets in a frame should be similar. For this reason, adjectives such as *blue* and *broken* which refer to completely different kinds of states, are not in the same frame, even though they are distributionally very similar otherwise. Admittedly, this diagnostic is easy to apply in some cases and hard in others.

  Note that it is not reliably possible to base frame distinctions mainly on selectional restrictions. One may, for instance, conclude that two different frames are involved with *blue* and *broken* given that, in the naive physics sense, the kinds of entities that have color are much more numerous and spread out over a greater part of an ontology than the entities that can be broken (in the concrete physical sense). However, within the Self-motion frame the verbs *crawl*, *flit*, *slither*, and *walk* have fairly different selectional restrictions, though they are grouped together in the same frame.

- The pre-specifications that the frame-evoking elements give to various frame elements will be similar. This heuristic was applied in the Mass_˜motion frame containing lexical units such as *crowd*.v, *flock*.v, *pour*.v, *stream*.v, *swarm*.v, *troop*.v, etc. These LUs are not simply treated as part of Self_motion since they all require that the moving entity be a mass theme, generally made up of many individuals.

  However, the desideratum that all the LUs in a frame share selectional restrictions on a frame element is often relaxed in practice, as it would result in very fine-grained frames that are impractical to manage. Consider, for instance, the requirement of the verb *tie* that the CONNECTOR be a relatively long, thin,

flexible object. Other verbs that refer to attaching one object to another (or, viewed symmetrically, both to each other) either make very different specifications (e.g. *glue*, *staple*) or none at all (e.g. *attach*).

Aside from factors that are used to group words into frames, there are those that never enter into consideration.

- At all times, we lump together those groups whose semantic differences are due to general constructions of the language:
  - passive
  - middle
  - tense/aspect constructions
  - composition with extra-thematic frame elements

  The list does **not** include causativity, since, as argued above, it is lexically specific.

- We group together words that are antonyms of each other. For instance, the adjectives *high* and *low* are both in the Position_on_a_scale frame. Similarly, the verbs *love* and *hate* are both in the Experiencer_subj frame. However, so-called converses (Cruse 1986), relational opposites such as *buy* and *sell* which take opposite points of view on a single event, are placed into separate frames because they profile different (sets of) participants. Likewise, so-called reversive pairs such as *tie* and *untie* are placed in different frames since the members of the pair denote different kinds of actions basically.

- Any differences purely in speech context.
  - deixis (*come* v. *go*)
  - register (*botch* v. *fuck up*)
  - dialect (*lorry* v. *truck*; *fixture* v. *regular season game*)
  - evaluation (*criticize* v. *praise*; *genius* v. *moron*)

  We mark these kinds of differences with semantic types (see Sect. 6.2). For example, the adjective *generous* in the Stinginess frame would be marked with the semantic type Positive_judgment, while the adjective *stingy* would be assigned the semantic type Negative_judgment. Note, however, that so far we have not applied semantic types in all cases where this would be appropriate.

Finally, frame development focuses on paraphrasability (or near-paraphrasability) of words and multiwords. That is to say, we ask whether one can more or less felicitously substitute one lexical unit for another and still evoke the same frame and express the same kinds of semantic roles as syntactic dependents of the new lexical unit. Frame development does **not** directly address the (near-)paraphrasability that may exist between whole utterances.

There are, for instance, many pairs of utterances where the total meaning of one member is built up compositionally from several frame evoking elements while the total meaning of the other derives just from a single lexical unit that combines within it the complex semantic structure that can be expressed with multiple frame evoking elements.

Clear cases of this situation include causative-inchoative pairs. Sentence (7) is centered on the inchoative *harden*, but combines it with the extra-thematic frame element *Cause*, expressed by the phrase *due to the hydration of the cement*, which composes to a causative scenario. Sentence (8) directly encodes the more complex causative scenario: there is no lexical material present that encodes the notion of causation separately from the notion Change_of_consistency.

(7) The paste **hardened** due to hydration of the cement.

(8) The hydration of the cement **hardened** the paste.

Note that although causative-inchoative pairs are not in the same frame, the FrameNet database provides an explicit link between the paired frames via a frame-to-frame relation **Causative_of**.

Similarly, sometimes the frame relation Using also connects frames whose targets can figure in utterances that are paraphrases of each other. Consider examples (9) and (10).

(9) I've **communicated** by email with US Airways on many occasions and have always gotten a reply within a few days.

(10) I've emailed US airways on many occasions and have always gotten a reply within a few days.

The lexical unit *communicate* in (9) belongs to the very general Communication frame, which is used or inherited by many frames. The lexical unit *e-mail* in (10) belongs to the Contacting frame, which uses the Communication frame. (For a more detailed discussion of the various frame relations, see section 6). The two lexical units and frames are not equivalent, which can be illustrated easily with the following pair of past tense utterances:

(11) I **e-mailed** him my new phone number.
(12) I **communicated** my new phone number to him by e-mail.

While (12) entails that the RECIPIENT received the message, (11) does not carry the same entailment. With Contacting, no actual successful communicative act is implied, only the successful completion of acts which could establish the communication.

Another kind of paraphrase between utterances relies on world knowledge and inference. Consider the pair of utterances in (13)-(14). Utterance (13) denotes a means action and utterance (14) denotes what is achieved by the means action.

(13) The Denver tight end caught the ball in the end zone.

(14) Denver scored a touch down.

Catching the ball in the end zone of the field is a means of scoring a touch down, but catching and scoring have no conventional relation to each other outside of football, and even in the football context the two are not necessarily linked.

However, in some cases when inference is needed to establish that sentences are paraphrases of each other, the lexical material may still guide the inferential process to some extent. Compare sentences (15), which denotes a means action, and (16), which denotes an accomplished GOAL.

(15) You needed my help so *I got on the bus.*

(16) You needed my help so *I came.*

Boarding a bus and arriving somewhere have no necessary relation to each other. However, the noun *bus* is a lexical unit in the Vehicle frame and that frame is linked to the Motion frame, which in turn is used by the Arriving frame that includes *come*. The framal links between *bus* and *come* thus provide some evidence that a semantic equivalence may be intended.

Generally, FrameNet groups words only for paraphrasability between lexical units but not for paraphrasability of utterances. There is, however, one kind of hybrid case where FrameNet groups lexical units in the same frame, even though the alternating lexical units do not participate in the same syntactic constructions and despite the difference in the relative prominence that the evoked frames have in the sentence. Consider the following pair of sentences:

(17) She **wisely** sold the house.

(18) She was **wise** to sell the house.

The understandings of (17) and (18) are clearly similar. However, there is a difference between the sentences in terms of the relative salience of the two frames, Mental_property and Commerce_sell. (17) mainly reports an event of selling whereas (18) focuses on the assessment of what the selling behavior says about the seller's mental abilities. The lie-test shows that this is so. If a speaker challenges (17) by saying *That's a lie*, the selling event is being contested, but if they challenge (18) in the same way, what's denied is the appropriateness of the assessment that it was *wise* to sell.

This pragmatic difference fails to be captured by many current logical representation formalisms of adverb meanings, which suggest that the adverb has the clause in its scope. While FrameNet's practice of grouping the adjective and adverb in the same frame is in line with the logical-semantic similarity, it does not match the pragmatic-syntactic fact that the frame evoked by the syntactic head of the clause is more prominent.

Chapter 3

FrameNet Annotation

3.1 Introduction

As a technical matter, the way in which FrameNet analyzes instances of a target predicate consists of marking up parallel aligned layers of annotation with appropriate label sets, as shown in Figure 3.1. The layers that are displayed in the FN Desktop can be manually selected by annotators. The number of layers and the kind of information that can be recorded on them is technically unlimited. But in FrameNet's current practice the four main annotation layers are the Target, frame element (FE), grammatical function (GF), and phrase type (PT) layers. On the first, the (parts of the) target predicate are marked while on the latter three, labels are applied to the constituents expressing the frame elements of the target.

The next-most important set of layers consists of the layers called Other; a layer called either Noun, Verb, Adj, Adv, or Prep depending on the part of speech of the target (this layer is also often called the part-of-speech-specific layer); and the Sent(sentence) layer. The Other layer holds labels relating to certain special constructional contexts in which the target may occur, such as relative clauses, existential constructions, and extraposition constructions. The part-of-speech-specific layer holds labels that can occur only with predicates of a particular part of speech. For instance, the COPULA does not occur with verbal targets; it can be applied on the Noun, Adj, and Prep layers of targets of the appropriate part of speech. The Sent(sentence) layer is special in that it does not actually bear any annotation labels: when the layer is invoked, information about the sentence as a whole can be recorded on an appearing list of check-boxes.

A final group of layers includes, among others, layers holding labels related to part of speech (POS) and Named Entity Recognition (NER). This information is derived automatically from our corpora and third-party software and is generally not modified by FrameNet annotators.[1]

We now turn to FrameNet's annotation process. The work can be divided into two kinds according to the way in which sentences are chosen for annotation. In the **lexicographic annotation** mode, our main focus is on the goal of recording the range of semantic and syntactic combinatory possibilities (valences) of each word in each of its senses. To that end, we extract sentences from the different texts of a corpus because they contain a predetermined target LU. Then we annotate a selection of the extracted sentences in respect to that particular LU.

In another kind of work, the annotation of running text (also called **full-text annotation**), the sentences are chosen for us, so to speak. Annotation of running text is technically possible thanks to the annotation layering technique: FN lexicographers can one by one declare each word in a sentence a target, select a frame relative to which the new target is to be annotated, get a new set of annotation layers (frame element, grammatical function, phrase type) and appropriate frame element tags, and then annotate the relevant constituents.

Before going further into the details of annotation, let us briefly consider the Revenge frame, which will figure as an example frame repeatedly in this chapter. The definition of this frame follows:

> An AVENGER performs some PUNISHMENT on an OFFENDER as a response to an earlier action, the INJURY, that was inflicted on an INJURED_PARTY. The AVENGER need not be identical to the

[1] Consequently, the default setting of the software is to not display this final set of layers but annotators can manually select them for display.

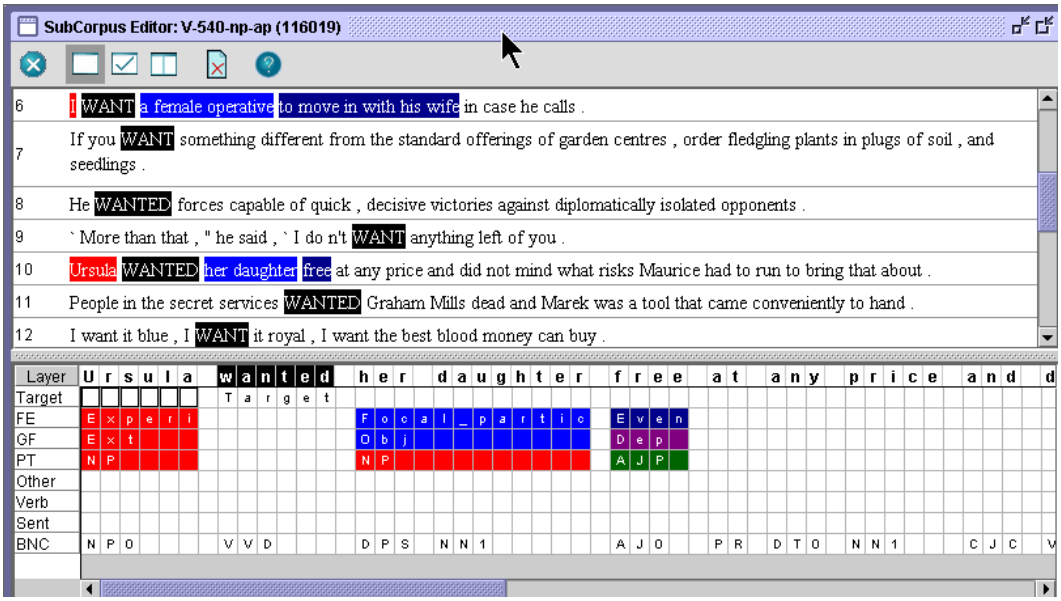

Figure 3.1: Annotation layers

INJURED_PARTY but needs to consider the prior action of the OFFENDER a wrong. Importantly, the punishment of the OFFENDER by the AVENGER is seen as justified by individual or group opinion rather than by law.

Thus, the frame elements in the Revenge frame are AVENGER, PUNISHMENT, OFFENDER, INJURY, and INJURED_PARTY.

Lexical units in this frame include *avenge.v, avenger.n, get even.v, retaliate.v, retaliation.n, retribution.n, retributive.a, retributory.a, revenge.v, revenge.n, revengeful.a, revenger.n, vengeance.n, vengeful.a*, and *vindictive.a*. Some example sentences with the lexical unit *avenge* are given here.

(1) [His brothers AVENGER] **avenged** [him INJURED_PARTY].
(2) With this, [El Cid AGENT] at once **avenged** [the death of his son INJURY].
(3) [Hook AVENGER] tries to **avenge** [himself INJURED_PARTY] [on Peter Pan OFFENDER] [by becoming a second and better father PUNISHMENT].

Regardless of the kind of annotation that is done, the following guidelines apply to the annotation relative to a particular instance of a target word.

- FrameNet annotation is directed toward the dependents of a target word. We do not annotate constituents that are understood only through context as referring to the filler of a particular frame element role. Consider the following example.

  (4) Smith was surprised when Lowry **retaliated** for the attack.

  The target *retaliate* evokes the Revenge frame introduced above, but the surface syntax does not contain constituents representing the OFFENDER or the INJURED_PARTY. One plausible understanding of the sentence as a whole is that Smith was the one against whom Lowry retaliated. One might, thus, want to tag *Smith* as the OFFENDER. However, there is an alternative interpretation: Smith could just be an observer of a conflict involving Lowry and a third party. In such cases, where there is no syntactic construction present that guarantees that interpretation, we do not annotate the non-local phrases in question. In the particular example here, we mark the OFFENDER as omitted under definite null instantiation (DNI).

- We annotate whole constituents rather than just the head words of the target's syntactic dependents. For instance, in example (3) above, the preposition *on* is included in the constituent expressing the OFFENDER frame element, *on Peter Pan*.

- Each dependent is annotated for Frame Element identity, Phrase Type, and Grammatical Function relative to the target LU.[2]

While the above guidelines apply equally to lexicographic and full-text annotation, there are also some clear differences between the two modes of annotation.

- In FrameNet's lexicographic work, annotation is done relative to only one lexical unit per sentence, the target. (As noted previously, the target is typically a single word but can also be a multi-word expression such as a phrasal verb (e.g., *give in*) or an idiom (e.g., *kick the bucket*).) In full-text annotation, by contrast, all contentful lexical units are treated as targets and their dependents are annotated.

- In full-text annotation, we have no choice but to annotate each instance of a target word whether or not it contributes useful lexicographic information about the target. For instance, in a given text the LU may occur in complicated syntactic constructions that add nothing to our understanding of the target's meaning. A particular use of the LU may be very poor style or borderline ungrammatical. Some instances only provide more exemplification of an already amply documented argument realization pattern for the LU rather than exhibiting a rarer pattern. This contrasts with our sentence-annotation (lexicographic) work, where it is intentionally part of our annotation principles to avoid such sentences.

- In the lexicographic work, we intend the set of sentences annotated for a given LU to represent the full range of combinatorial possibilities for that LU. A given text, even of considerable length, is not likely to exemplify all patterns in which a target may occur.

- The annotation contained in the FrameNet database does not offer information about frequency of occurrence. Full-text annotation necessarily means that we annotate all instances of any given contentful target. Thus, at least relative to a particular text, frequency information for LUs can be derived.

So far, we have sub-divided our work into two parts depending on whether we choose the sentences to annotate or whether they are chosen for us by a text. Another subdivision applies specifically to our lexicographic work where we produce annotation of two different types, reflecting two different kinds of target words:

- annotation relative to a **frame bearing** syntactic governor, either a predicate, modifier, or a referring expression and

- annotation relative to a **slot filler**, that is, relative to a referring expression that is a frame element of a frame determined not by itself but by a governor

Both kinds of annotation will be discussed. Since the annotation relative to frame-bearing syntactic governors is FrameNet's main task, it will be discussed first and in greater detail in sections 3.2-3.7. Annotation relative to slot-fillers is discussed in section 3.8.

3.2 Lexicographically motivated annotation practices

In accord with FrameNet's goals, syntactic and semantic descriptions are tailored to lexicographic description only and may differ from regular linguistic analysis as well as from shallow semantic analysis in several ways.

- Typically, the analysis of sentences is only **partial** in the sense that we apply labels just to some words or phrases of a sentence. Although many sentences contain multiple frames evoking lexical units, in normal annotation we annotate only with respect to one of them. Also, there are some classes of frame evoking elements that we have not yet tackled, an important one being sentence connectives such as *although, while, but,* etc.

[2] As will be discussed in section 3.2, we also sometimes apply frame element labels on a secondary frame element annotation layer that are not paired with Phrase Type and Grammatical Function labels. These annotations do not contradict the statement concerning the annotation of dependents with triples of FE, PT, and GF since they involve words or phrases that are not themselves direct dependents of the target.

- We tag **whole constituents** that realize frame elements relative to our target words, rather than just tagging the head words of these constituents. That is, we work with a phrase structure grammar, rather than a dependency grammar.

  A consequence of this is that many frame element labels cover words that have no direct relation of their own to the target, but only to the head of their constituent. For instance, when a frame element is expressed by a noun which takes adjectival, prepositional or clausal complements or which is modified by such elements, these complements and modifiers are included in the frame element tag. Consequently, phrase type (and grammatical function) labels always cover full constituents. In (5), we tag both the noun *story* and the whole NP it heads; in (6), we tag both the noun *fact* and the full NP it heads; in (7), the *with*-PP is part of the frame element that is tagged.

  (5) I **heard** [a story *about a man named Jed* NP].
  (6) [The fact *that I wear a funny shirt* Sfin] is totally **irrelevant**.
  (7) I **have** [a cat *with orange stripes* NP].

- FrameNet syntax also differs from traditional treatments of subcategorization in paying attention to **relational modifiers**, including the non-heads of compounds. For example, the FrameNet description of a noun like *clinic* makes reference to the types of nouns which can modify this noun in compounds like *allergy clinic*. In theoretical treatments, modifiers such as *allergy* relative to *clinic* are typically assumed to be outside the realm of subcategorization. FrameNet includes them because they frequently express the same semantic roles (or frame elements) with respect to the modified heads as complements express with respect to their governors. For example, the Cure frame associated with the verb *treat* includes a role for the Affliction which is treated, and this role is typically expressed as the object of *treat*. Thus, you can *treat an allergy*, *treat the flu*, and so forth. Modifiers occurring with the noun *treatment* often express the same role; thus, there are *allergy treatments*, *flu treatments*, etc. Such observations are of lexicographic relevance, and our syntactic descriptions reflect this perspective.

- In some situations, there are differences between the syntactic and the semantic **headedness** of a clause. For instance, *Pa gave her a lecture* has *give* as its syntactic head. However, from a semantic point of view the sentence reports an act of lecturing, not one of giving. We understand cases such as *give a lecture* to involve frame-evoking nouns that are syntactically 'supported' by verbs in order to be able to project clauses. This analysis is strengthened by the fact that noun-support verb combinations typically involve the selection of the verb by the frame-bearing noun (**make/give a lecture* v. *make/*give an attempt*). An important consequence of this analysis is the annotation of subjects of support verbs as frame elements relative to the noun. Thus, the example sentence *Pa gave her a lecture* would be annotated with respect to the target *lecture*, with *gave* marked as a support verb, *Pa* marked as the Speaker and *her* marked as the Addressee.

- Syntacticians commonly divide the complements of a predicate into arguments and adjuncts, or arguments and modifiers. In FrameNet, the semantic spirit of this distinction is covered by the distinctions of **coreness** status rather than through GFs (cf. section 3.2.1). As a consequence, we only have a grammatical function Dep (dependent), rather than distinguishing between the two types of complements.

- In general, we select sentences for annotation where, with the exception of subjects, all frame elements are realized locally by constituents that are part of the maximal phrase headed by the target word. In addition, there are several clear cases, typically involving subjects, where the combination of particular lexical items and grammatical structures guarantees a particular interpretation of phrases which are not dominated by the maximal phrase headed by the target. One such example is that of modal verbs, which behave as so-called raising verbs. The subjects of modals such as *may*, *will* etc. always need to be interpreted as subjects and frame elements of the predicates that follow the modal. For an example like *[Bill SELLER] may **sell** it at that price*, a treatment motivated by syntactic theory might suggest that the subject slot of *sell* is filled by a kind of trace, or that one of its required arguments is passed on to the higher raising predicate via a valence passing principle. Under such a treatment, it could be left up to grammatical principles to establish the connection between the subject of the raising verb and the valence requirement of the target verb. However, from a lexicographic point of view it is more useful to directly report what the semantic properties of the fillers of particular Frame Elements are, rather than to ask users to employ additional tools to establish lexical antecedents or valence fillers.

- For most targets, frame elements are marked only on a single frame element layer. Under two kinds of circumstances, however, we record frame elements on an additional, secondary frame element layer.[3]

  The first type of situation for which second-layer annotation is used is **frame element conflation**, when a phrase that directly expresses one frame element also allows inferences about a second frame element. Consider as an example the sentence *Doctor Miller managed to at least **cure** [the epileptic]*. Out of context, this sentence is most naturally understood to mean that the epileptic is cured of epilepsy. We tag the phrase *the epileptic* as PATIENT on the first frame element annotation layer and as AFFLICTION on the second annotation layer. Notice that when a core frame element, such as AFFLICTION in the example of *cure*, is applied on a secondary annotation layer (cf. 3.2.5), no null instantiation tag is used concurrently.

  Given an appropriate context, all inferences in conflation contexts can be defeated. In cases where the inference is defeated, no secondary layer tagging is applied. In the modified sentence *Both the epileptic and the schizophrenic patient had terrible head colds last week and Doctor Miller managed to at least **cure** [the epileptic]*, the AFFLICTION frame element is marked as null-instantiated rather than applied to the constituent *the epileptic* on a second frame element layer. Note that frame element conflation is the only context where FN annotation marks phrases that are only inferred, rather than grammatically guaranteed, to express information about frame elements.

  A second type of circumstance where FN annotation makes use of a secondary annotation layer has to do with **possessor phrases** that are obligatorily co-referential with another argument. Most commonly this arises with body part possessors such as in the Experience_bodily_harm frame. With verbs such as *stub*, *sprain* or *strain* the referent of the EXPERIENCER frame element has to be expressed again as a possessive determiner in the BODY_PART frame element and we tag the possessive on the second frame element layer with another instance of the frame element EXPERIENCER. Thus, in *Ike **cut** [his finger] on the envelope*, the subject NP *Ike* is marked as EXPERIENCER on the first annotation layer, the phrase *his finger* is marked as BODY_PART, and the possessive *his* is marked as EXPERIENCER on the second frame element layer.

  In all cases, a label applied on the second layer is always co-extensive with, or contained in the boundaries of a label on the first FE layer; no word or constituent can be covered by a frame element label on the second layer alone. Finally, secondary frame element layers are never accompanied by grammatical function or phrase type layers; GF and PT layers are available only for first layer annotation.

- The combination of our surface-oriented annotation and our commitment to providing grammatical function and phrase type information about the phrases that fill frame element roles causes some problems in the case of **discontinuous frame elements**. For instance, with *wh*-extraction out of a frame element such as *[Who] did you **talk** [about]?*, we need to provide a grammatical function tag for the fronted *wh*-word *who* even though it does not bear a grammatical relation to the target word *talk*. This and other cases of discontinuous frame element realization are discussed below in section 3.2.8.

3.2.1 Coreness

We classify frame elements in terms of how central they are to a particular frame, distinguishing three levels: **core**, **peripheral**, and **extra-thematic**. A fourth possible value for this attribute, called **core-unexpressed** is also discussed below.

A core frame element is one that instantiates a conceptually necessary component of a frame, while making the frame unique and different from other frames. For example, in the Revenge frame, AVENGER, PUNISHMENT, OFFENDER, INJURY, and INJURED_PARTY are all core frame elements, because an avenging event necessarily includes these participants. One cannot imagine an act of revenge that is not preceded by a (perceived) offense or one that is not directed against anybody.

In determining which frame elements are considered core, we also consider some formal properties that provide evidence for core status. These properties are typically co-present, although they need not be.

- When an element always has to be overtly specified, it is core. For instance, the verb *resemble* in the Similarity frame always requires a post-verbal complement NP denoting an entity that is similar to the entity denoted by the subject.

[3]Technically, it is possible to have more than one additional frame element layer but this capability is almost never used in practice.

- A frame element which, when omitted, receives a definite interpretation, is also core. For instance, when the verb *arrive* is used only with a THEME-subject, as in *John arrived*, a particular GOAL location that the THEME reaches still has to be understood in the context. GOAL, therefore, must be a core frame element.

- A frame element whose semantics cannot be predicted from its form, in particular from any marking prepositions, ought to be core since its interpretation completely depends on the target. From this, we can derive two corollaries, given below.

  – A frame element that has no formal marking should be core. Thus, frame elements that can be subject or object in a simple active sentence ought to be core since these slots host frame elements of many different kinds and knowing that something is a subject or object is not particularly informative. For example, the Building frame which contains the verb *build* has at least two core frame elements since the verb *build* has both a subject and an object.

  – A frame element that has idiosyncratic formal marking should also be core. A good example of this is the prepositional verb *depend **on***. The preposition *on* does not occur as a marker of the same meaning with predicates in many other frames. In its basic spatial sense of 'in contact with and supported by', *on* occurs in many different frames; as a marker of PLACE or LOCATION frame elements it is totally unremarkable and does not suggest core status for these FEs.[4]

Frame elements that do not introduce additional, independent or distinct events from the main reported event are characterized as peripheral. Peripheral FEs mark such notions as TIME, PLACE, MANNER, MEANS, DEGREE, and the like. They do not uniquely characterize a frame, and can be instantiated in any semantically appropriate frame. In respect to the Revenge frame, any report of an event of revenge may also include explicit information about the parameters of time, place, manner, etc. of the revenge, an example of which is given below.

(8) The bereaved family **retaliated** [immediately TIME].

Extra-thematic frame elements situate an event against a backdrop of another state of affairs, either of an actual event or state of the same type, as illustrated with ITERATION, or by evoking a larger frame within which the reported state of affairs is embedded, as shown for CONTAINING_EVENT.[5]

(9) Thou shalt not exact **revenge** [twice ITERATION] for the same offense.
(10) The Aussies took **revenge** [in a penalty shootout before 2465 fans in Long Beach the next day Containing_event].

Note that extra-thematic frame elements are understood not to conceptually belong to the frames they appear in. We take them to properly be frame elements of other abstract frames that take them as well as the targets that they modify as arguments. Thus, in example (11), we take *twice* and the verb phrase *eat* to be arguments of a more abstract Iteration frame. Similarly, in example (12), *cooked dinner* and *me* are frame elements of a Benefaction frame. Note that, as shown by (12), the native frame of the extra-thematic frame element need not be evoked by lexical material, it may simply be evoked constructionally.

(11) Learn how to spend a few extra minutes planning complementary menus where you cook once and **eat** [twice].

(12) Lennert, another sweetie in my life, **cooked** [me] dinner, mmm mmm good.

The view of extra-thematic frame elements presented here entails that these frame elements are necessarily the same across all the 'host' frames in which they appear. That is, unlike core and peripheral frame elements, extra-thematic frame elements do not have a frame-specific understanding. By comparison, although many core frame elements named AGENT share properties with each other due to Inheritance and Using relations,

[4]This second corollary also supports the argument made in Chapter 2 that causative and inchoative frames should be separated. When CAUSE-like FEs appear in inchoative sentences they are marked by prepositions such as *because of*, *due to* that carry the right causal semantics all by themselves. The Cause-like FEs are thus not solely dependent for their interpretation on the target lexical unit.

[5]The frame hierarchy that is under construction offers another way to think about the status of extra-thematic frame elements. We must assume that, unlike the core or peripheral frame elements of a given frame, the extra-thematic frame elements occurring in that frame do not occur as core frame elements in any of the ancestor frames, that is, frames to which the current frame is related by inheritance.

they do not *necessarily* have identical properties. More importantly, one cannot predict the frame role AGENT from the fact that an NP has the grammatical function Ext.

(For an overview of the most frequently occurring extra-thematic frame elements, the reader is referred to Appendix A.)

The value "Core-Unexpressed" is a special notational shorthand. It is assigned to FEs that behave like core frame elements in the frame where they are marked as Core-unexpressed but which, counter to expectation, may not be used for annotation in descendants of that frame. Frame elements marked as core-unexpressed will thus not necessarily be listed among the FEs in descendant frames.

We do not consider core-unexpressed frame elements to be violations of our definition of full inheritance. Our reasoning can be illustrated with the example of the core-unexpressed ACT frame element in the Intentionally_act frame, which is exemplified in (13).

(13) I'll **do** [the vacuuming ACT].

In the many child frames of Intentionally_act such as Choosing, Perception_active, etc., the idea of an ACT is as relevant as in the Intentionally_act parent frame. However, in the child frames the frame element is absorbed by the lexical units in the frame and cannot be separately expressed.

Marking the frame element ACT as core-unexpressed in the Intentionally_act parent frame allows us to keep the frames that are lower in the hierarchy from including an inherited FE which for any lexical unit in the frame could at most be annotated on the target itself, but never be expressed separately. The sentence **I **chose** decision the blue one* is simply ungrammatical.

Coreness marking makes the most sense for event and state frames, and in these frames we use all three statuses. Coreness marking is done at the level of the frame and is intended to be consistent for all lexical units in a frame. In frames whose LUs are artifacts or natural kinds, we only use a two-way distinction (core and peripheral), noting that the values do not have exactly the same meaning as with events. In such cases, typically there is just one core frame element which is marked on the target word. For instance, in the Clothing frame the FE CLOTHING is core, and all other FEs are peripheral.

3.2.2 Frame element relations

In providing a semantic analysis of the combinatorial possibilities of our target predicates–rather than one stated only in terms of phrase types and grammatical relations–we have observed that in an important sense frame elements are not independent of each other. Frame elements are related to the frame and required by it, as well as interrelated directly in a number of ways. These interrelationships have a direct impact on annotation, as they license the absence of core frame elements, which must normally all be accounted for in every annotation set, or (more rarely) require frame elements that might otherwise be optional. FrameNet systematically records these interrelationships. The three types of frame element relations that we recognize are discussed in sections 3.2.2.1-3.2.2.3. Section 3.2.2.4 discusses some respects in which the current treatment of frame element sets needs further refinement.

3.2.2.1 Coreness Sets

In our annotation practice, we often find that some groups of FEs seem to act like sets, in that the presence of any member of the set is sufficient to satisfy a semantic valence of the predicator. We refer to such a group of FEs as a **coreness set**, or **CoreSet**. For instance, SOURCE, PATH, and GOAL are core FEs in the various motion frames in the database. However, although possible, it is not necessary, and in fact unusual, for all three FEs to co-occur, as in example (14). Sentences in many motion frames can be informationally complete and pragmatically felicitous with just one or two of the FEs expressed, as shown in (15)-(18).

(14) Fred **went** [from Berkeley SOURCE] [across North America and the Atlantic Ocean PATH] [to Paris GOAL].
(15) Martha **hiked** [from Berkeley SOURCE] [to Oakland GOAL].
(16) Elaine **walked** [to Monterey GOAL].
(17) I saw Peter **sneak** [past the guard PATH].
(18) Juan was **walking** [out of the office SOURCE] when I arrived.

FrameNet's normal annotation practice demands that we account for all core FEs and we could keep track of the un-instantiated FEs in the example sentences above by using null-instantiation labels as described in section (3.2.3). However, we prefer to group the FEs in CoreSets and not mark null instantiation for each

member FE in cases where the FEs have an informational and conceptual interdependence. SOURCE, PATH, and GOAL, for instance, are clearly related via a notion that we might call 'full path'. By contrast, omission of the INGESTIBLES for *eat* in the Ingestion frame is not related to the presence or absence of any other frame element. The sentence *I'm **eating** __ (with my friends/in the kitchen/now)* is acceptable with or without any of the frame elements given in parentheses expressed.

3.2.2.2 Requires

In some cases, the occurrence of one core FE requires that another core FE occur as well. To illustrate, in the Attaching frame ITEM, GOAL, and ITEMS all are core FEs. If ITEM occurs, then GOAL is required, as shown below, where the sentence without a GOAL is unacceptable. In this situation, we mark a Requires relation between the two frame elements.

(19) The robbers **tied** [Paul ITEM] [to his chair GOAL].
(20) * The robbers **tied** [Paul ITEM].

The Requires relation occurs in almost all frames that have a construal alternation between a symmetric construal, when a single frame element name is used, and an asymmetric construal, when two frame elements with names of the are used. In the former case a simple name of the form [FENAME]s is used and in the latter, two FEs of the form [FENAME]_1 and [FENAME]_2 are used. Some sample frames are Compatibility, Chatting, and Similarity; there are many more.

3.2.2.3 Excludes

In some cases, if one of the FEs in a group of conceptually related FEs shows up, no other FE from that group can. Again, in the Attaching frame, if ITEMS occurs, then ITEM and GOAL are excluded. In this situation, we say that ITEMS excludes ITEM and GOAL.

(21) The robbers **tied** [his ankles ITEMS] together.

The above Excludes-relation in the Attaching frame is an instance of a much more common pattern of alternation between a symmetric/reciprocal and an asymmetric construal of events or states involving two parties. In most frames, where the alternation is possible, the names of the frame elements reflect the underlying alternation between reciprocal and asymmetric construal. For instance, in the Similarity frame with lexical units such as *similar*, *different*, etc. we have the frame elements ENTITY_1 and ENTITY_2, and ENTITIES. Usually, one is allowed to infer equal participation in the event or state by the grammatically less profiled participant (PARTICIPANT_2). However, since in the case of Attaching, the GOAL (which would be ITEM_2 under our normal naming scheme) is not readily understood as itself being secured or immobilized via attachment to the ITEM on the asymmetric construal, we selected a name that reflects the fact that the usual inference to equal status for PARTICIPANT_2 in the asymmetric construal is not warranted.

The Excludes relation also manifests in frames where an event can be brought about either by an intentional AGENT or by a CAUSE event. Consider the following examples from the Placing frame.

(22) [The same flood tide that had brought such a good harvest of tiles CAUSE] **heaped** a mass of driftwood onto the Reach.

(23) [Bill AGENT] **deposited** the bag of croissants and the Financial Times carelessly on the hall table.

The two sentences represent two different construals of Placing scenes. Sentence (22) focuses on an event as causing the change in location of the THEME, whereas sentence (23) focuses on an AGENT who through their involvement in an unspecified event, most likely an intentional action involving his hands and body, causes the change of location of the THEME. The two construals are incompatible (since there is only one subject slot) and the frame elements AGENT and CAUSE stand in an Excludes relation to each other.[6]

[6]In any sentence in which the event that the AGENT is involved in is specified together with the AGENT, the AGENT would have to be realized as subject and the event as a clause or PP. In such sentences, we would call the event a MEANS rather than a CAUSE.

Another clear instance of the Excludes relation between frame elements occurs in the Evading frame, where an EVADER moves under its own power to thereby avoid CAPTURE or contact with a PURSUER. The CAPTURE is an actual or hypothetical event in which the PURSUER takes physical control of the EVADER. The CAPTURE frame element and the PURSUER are thus clearly interrelated but only one of them can appear as a dependent of a target in the Evading frame.

(24) Sheriff's officials said they apprehended a gang member after he **evaded** [them PURSUER].

(25) He had successfully **evaded** [arrest CAPTURE].

Finally, note that the Excludes relation strictly applies only to the direct syntactic dependents of a target word, that is, to first layer annotation. Frame elements that exclude each other may co-occur in an annotation set if they appear on separate annotation layers.

(26) Perkins McLain **evaded** [capture [through Spain PURSUER] CAPTURE] .

(27) The **discussions** [between [Miller INTERLOCUTOR_1] and [the dean INTERLOCUTOR_2] INTERLOCUTORS] went nowhere.

In (26), information about the PURSUER is expressed inside the CAPTURE frame element in a prepositional phrase dependent of the noun *capture*. In (27), the two sides of the discussion, INTERLOCUTOR_1 and INTERLOCUTOR_2 are expressed within the coordinate NP that encodes the INTERLOCUTORS frame element.

3.2.2.4 Future refinements

The treatment of frame element relations sketched in the preceding sections is adequate for a large number of frames. However, two systematic problems remain.

One problem is that we have no explicit treatment of the idea of **proto-frame elements**, of which other frame elements are more specialized expressions. In cases like (28), we would prefer not to have to pick either specifically AGENT or CAUSE as constructionally null instantiated, since the context might not provide enough information to resolve that question. Instead we would make reference to a superordinate frame element (call it *FORCE) that is vague about intentionality and the event-person distinction. Likewise in (29), where B answers a question about a new employee, we would prefer to use a superordinate frame element (which might be named *FIELD in this case), rather than choosing among the frame elements ROLE, SKILL, KNOWLEDGE, or FOCAL_PARTICIPANT for constructional null instantiation relative to the predicate *good* in the Expertise frame.

(28) The car got **damaged** while parked outside of our house.

(29) A. How's Susan working out? B. She's very **good**.

Having an explicit representation of proto-frame elements would also be useful in dealing with certain linguistic expressions that seem to instantiate the superordinate proto-frame element rather than one of its more specific manifestations. Consider the phrase *trench to trench* in (30): it does not refer to the SOURCE or GOAL of a PATH but neither does it refer to a middle PATH which would be compatible with the specification of a final GOAL. A proto-frame element FULL_PATH would provide the most adequate treatment.

(30) He **crawled** [trench to trench], looking for some sign of Stephen.

Another use for Proto-frame elements involves inheritance relations. In some cases, an inheriting frame will allow only one FE from an Excludes or CoreSet group in the parent frame. Superficially, this violates the rule that requires child frames to have a corresponding FE for each core/peripheral FE of the parent. In a deeper sense, however, inheriting only one member of a frame element set should be permitted on the understanding of inheritance as subtyping. This is so because the child frame is fulfilling every constraint of the parent, merely adding a constraint that prevents one of the construals possible in the more generic case. If we state the frame element restriction on inheritance so that it pays attention only to the most generic level of FEs, then mappings from subsidiary FEs are allowable, but not required. This would make it possible, for instance, to have a *Murder frame (with only agentive causes) as a child frame of the Killing frame (which allows CAUSEs or AGENTs). [7]

[7] Currently, predicates denoting intentional killing are in the Killing frame together with predicates that allow either intentional or non-intentional causation.

The second major problem that remains concerns the treatment of **subject selection constructions**. The current treatment of coreness requires that all frame elements that can occupy the subject position be marked as core frame elements. In many frames this leads to MEANS and INSTRUMENT frame elements having core status and further being part of a CoreSet with AGENT, since the two frame elements may co-occur.

Consider the verb *open* in the Closure frame. In (31), we have a canonical AGENT subject but in (32) an INSTRUMENT fills the subject slot. (Note that the frame conceptually requires an AGENT; uses of *open* involving a CAUSE such as *The wind opened the door* are handled in another frame.)

(31) [John AGENT] **opened** the door.
(32) [The key Bill gave him INSTRUMENT] **opened** the door right away.

Given the earlier discussion of Frame Development (in section 2), INSTRUMENT should be core in the Closure frame since the noun phrase realizing it in (32) appears in the subject position. However, were there a separate way of representing the fact that English allows frame elements that are situated between the end-points of a causal chain to occur as subjects, it would not be necessary to give INSTRUMENT core status. This would avoid introducing a coreness set of AGENT and INSTRUMENT.

A separate treatment of these subject selection facts would be parsimonious and would also expose the essential lexical similarity between the English frames and the frames of languages such as Japanese, where subject selection is much more restricted and INSTRUMENTs and MEANS rarely, if ever, appear as subjects.

3.2.3 Null instantiation

Sometimes FEs that are conceptually salient do not show up as lexical or phrasal material in the sentence chosen for annotation. Nevertheless, we indicate their absence since it provides lexicographically relevant information regarding **omissibility conditions**. The FE that has been identified indicates which semantic role the missing element would fill, if it were present.

With respect to null instantiation, verbal, adjectival, and prepositional targets are treated identically. (For null instantiation with noun targets, see 3.4.3.) The following examples show omitted elements with each part of speech. (The name of the frame element is given in square brackets and the frame of the target is given in parentheses.)

(33) That will **suffice** . [DNI STANDARD] (Sufficiency)
(34) The result should be **similar** . [DNI ENTITY_2] (Similarity)
(35) I tried to put the toys back **in** . [DNI GROUND] (Locative_relation)

Not all cases of frame element omission are alike. We recognize three different cases, one that is not lexically specific and two that are. We will now discuss the three types of omission, focusing on null instantiation with verbs where the phenomenon is clearest.

3.2.3.1 Definite Null Instantiation (DNI):

The first type of lexically specific null instantiation to be considered is the definite (or *anaphoric*) type. Cases of definite null instantiation are those in which the missing element must be something that is already understood in the linguistic or discourse context. In the following example, the OFFENDER is not expressed overtly in the syntax, but its referent has to be known in the context.

(36) [The monkey AVENGER] **avenged** [himself INJURED_PARTY] [by growing to the size of a giant and setting fire to the city PUNISHMENT]. [OFFENDER DNI]

3.2.3.2 Indefinite Null Instantiation (INI):

The indefinite cases (sometimes also referred to as *existential*) are illustrated by the missing objects of verbs like *eat*, *sew*, *bake*, *drink*, etc., that is, cases in which these ordinarily transitive verbs can be spoken of as used intransitively. (e.g. *Molly rarely eats alone*; *Granny begins baking about a month before Christmas*; *Michael even drinks heavily on weeknights*.) As is well known, there are often special interpretations of the existentially understood missing objects. For example, with *eat* the missing entity is likely to be a meal, with *bake* it is likely to be flour-based foods, with *drink* it is likely to be alcoholic beverages, etc. In contrast to anaphoric omissions, with existential omissions, the nature (or at least the semantic type) of the missing

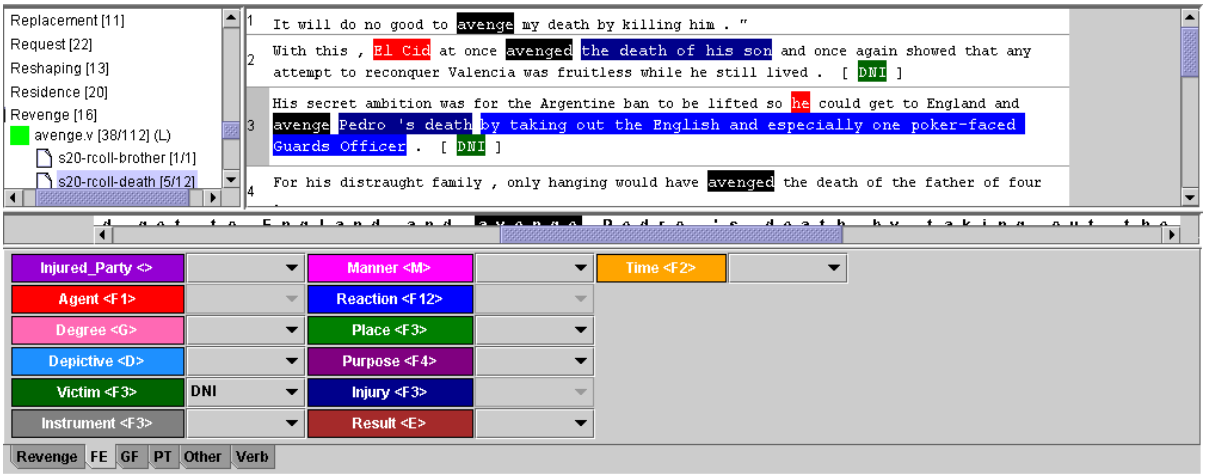

Figure 3.2: Annotating a Null Instantiated Frame Element

element can be understood given conventions of interpretation, but there is no need to retrieve or construct a specific discourse referent.

For example, in the Revenge frame, all lexical units happen to allow the frame element PUNISHMENT to be omitted under indefinite null instantiation. This is shown for *avenge* in (37).

(37) He took it out on Scarlet in the same way as [he AVENGER] **avenged** [himself INJURED_PARTY] [on her OFFENDER] [for the pressures at work and the demands of his first wife INJURY] . [INI PUNISHMENT]

Note that both in the case of definite and indefinite null instantiation, the LUs in a frame may differ from each other in whether or not they allow the omission. For instance, while *eat* allows its object to be omitted, *devour* does not, even though they are both in the Ingestion frame.

Verbs that usually require an argument to be present (or only allow it to be omitted under conditions of definiteness) can be used in a generic construction with indefinite null instantiation, as shown below.

(38) He **takes** and never gives back. [INI THEME]

Figure 3.2 shows the FrameNet Desktop opened for the annotation of a case of definite null instantiation in the sentence beginning *His secret ambition. . . .* Notice the tab to the right of the frame element VICTIM in the picture, which shows that the FE was omitted under definite null instantiation. In addition, an appropriately colored DNI tag appears at the end of the sentence in the corpus viewing section of the FNDesktop.[8]

3.2.3.3 Constructional Null Instantiation (CNI):

Constructionally omitted constituents (also called *structurally* omitted) have their omission licensed by a grammatical construction in which the target word appears, and are therefore more or less independent of the LU. Cases of CNI include: the omitted subject of imperative sentences, the omitted agent of passive sentences, the omitted subjects of independent gerunds and infinitives (i.e., the PRO-elements of generative grammar), and so on. In each of the following two examples, the FE AVENGER is tagged with the symbol CNI.

(39) Family feuds last for generations, and [slurs on honor INJURY] are **avenged** [by murder PUNISHMENT] . [CNI AVENGER]

(40) **Get even** [with her OFFENDER] [for this INJURY] [CNI AVENGER]

[8]In fact, the name of the INJURED_PARTY (the FE is shown as VICTIM in the figure, but has been renamed to INJURED_PARTY), Pedro, forms part of the FE INJURY; see Sec.3.2.5 below.

In addition, we use CNI for missing objects in instructional imperatives such as exemplified below, even though the omission is not dependent on a particular construction, but rather on particular genres, such as cookbooks and product labels.

(41) **Cook** on low heat until done. [CNI FOOD]
(42) **Tie** together loosely. [CNI ITEMS]

The experiential perfect also licenses object omissions that are not possible in simple assertions of frame instances.

(43) Have you ever **fostered** [CNI CHILD] before?

Note that particular constructions licensing argument omission specify particular interpretations, either indefinite (existential) or definite (anaphoric). For instance, the instructional imperative construction in (41) and (42) specifies a definite interpretation, whereas the experiential perfect exemplified in (43) specifies an indefinite interpretation. The CNI label thus collapses the interpretational distinction that we make among the lexically licensed omissions; it does not represent a separate kind of interpretation in addition to the definite and indefinite types.

3.2.4 Incorporation

There are many frames containing verbs which generally involve a particular of frame element, but where some of the verbs **incorporate** information about that frame element in their definition. Consider, for example, verbs of body movement; typically certain verbs are expected to co-occur with the name of a body part, even when the identity of the body part is clear from the meaning of the verb. A dog *wags* its tail, people *arch* their brows, *bat* their eyes, *purse* their lips, etc. But in the case of *smile, grimace, frown, pout*, and *scowl*, the affected body part is not separately expressed; we say that it is *incorporated*. Some verbs in this frame can optionally express the expected body part: one can say either *She blinked* or *She blinked her eyes*. Likewise, in the Placing frame, many verbs incorporate the GOAL FE (i.e. the place where the THEME ends up) such as *bag*.v, *bin*.v, *bottle*.v, *box*.v, *cage*.v, *crate*.v, *file*.v, *garage*.v. Note that it is still possible to further specify the incorporated FE explicitly, as in *They bottled the wine in custom-made blue bottles*, in which case, *in custom-made blue bottles* is annotated as the GOAL FE, as usual. In defining a new LU, one can specify an incorporated FE; the information as to what FE is incorporated in each LU is recorded in both the frame XML file and the LU XML file, and is displayed in the lexical entry report.

3.2.5 Frame element conflation

In some cases, information about two frame elements is expressed in a single constituent, a situation we call **conflation**. For instance, the concept of ousting somebody from office requires an understanding of the incumbent of the office and the identity of the office; each can be represented separately in a sentence like *We ousted Jones as mayor*. But in *We ousted the mayor*, the direct object stands for both the office and the incumbent.

We also find examples of frame element conflation in the Revenge frame. In particular, the INJURED_PARTY may be contained as a possessive in the phrase that realizes the INJURY frame element, as seen in the following example:

(44) [He AVENGER] **avenged** [Pedro's death INJURY] [by taking out the poker-faced Guards Officer PUNISHMENT].

Here, the possessive *Pedro's* realizes the frame element INJURED_PARTY, the person who suffered the INJURY. In such cases, the annotation tool allows for the creation of an additional FE layer, enabling the secondary annotation of (parts of) constituents in the same frame, as shown in Figure 3.3.

Note that there is never a phrase type or grammatical function indicated for the frame elements on the secondary FE layer.

Even with conflation, it is still possible for the FE annotated on the secondary layer to be instantiated as a constituent of its own in the syntax. In example (45), information about the person who was hit is encoded by the direct object *me*, as well as by the possessive determiner *my* inside the NP complement of the preposition *on*.

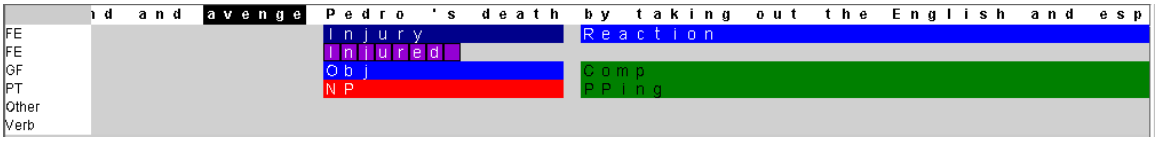

Figure 3.3: Secondary FE Annotation

(45) He **hit** me on my hand.

In such cases, we tag the possessive on the second layer with the same frame element label that is applied on the first annotation layer to the object of the verb.

3.2.6 Syntactic locality

In general, we select sentences for annotation where, with the exception of subjects, we find all frame elements realized by constituents that are part of the maximal phrase headed by the target word. There are two types of situations in which we annotate non-local constituents with frame element labels. In each case the motivation for annotating constituents that bear no syntactic relation to the target is lexicographic: the non-local constituents contain lexical material and as such are of interest to the study of collocations since they provide more information about the semantic type of the frame element than the locally occurring co-indexed phrases or empty elements do.

The first case in which we annotate non-local constituents consists of cases in which the target word is syntactically governed by a raising or control predicate. In such cases, the valence properties of the higher raising or control predicate guarantee that one of its arguments is also interpreted as an argument of the target, even though the relevant argument is not dominated by the maximal phrase headed by the target. Some of the most common types of control and raising predicates are illustrated below. The control or raising predicate that guarantees the interpretation of the non-local noun phrase as a frame element of the target appears in `typewriter` font.

Raising

Subject to object

(46) We `expect` [John AVENGER] to **retaliate** [against us OFFENDER] [INI PUNISHMENT] [DNI INJURY].

Subject to subject

(47) [John AVENGER] `seems` to have **avenged** [the death of his brother INJURY] [by luring Smithers into a trap PUNISHMENT].

Control

Subject control

(48) [They AVENGER] are `hoping` to **get even** [with Smithers OFFENDER] [for the insult INJURY].

Object control

(49) The commander `ordered` [the troops AVENGER] not to **retaliate** [against the rebels OFFENDER]

Tough-movement

(50) [The defeat INJURY] was `difficult` to **avenge**. [CNI AVENGER]

Note that Raising and Control cases are not restricted to verbal controllers, i.e. nouns may also serve that function, as illustrated here.

(51) Only a short few weeks ago, even [my PERCEIVER] `hope` of **seeing** [her PHENOMENON] was just a dream [Subject control].

(52) Meanwhile, today, Americans, hungering for victory, are puzzling over the Pentagon's `order` to [the troops AGENT]not to **put** [the Stars and Stripes THEME] [on their vehicles GOAL]. (Object Control)

(53) The testers gave it a clear thumbs up both for [its CREATED_ENTITY] `ease` of **assembly** and sail performance. [Tough movement]

The second case in which we annotate clearly non-local constituents with frame element labels concerns targets that occur **inside** relative clauses. Here our policy is to tag not only the constituent containing the relativizer (if there is one) as a frame element but to also repeat the FE/GF/PT triple on the antecedent, and to further mark the relative word and the antecedent phrase on the Other layer. Thus, our annotation for the simplest cases with an overt relative word is as shown in Figures 3.4 and 3.5.

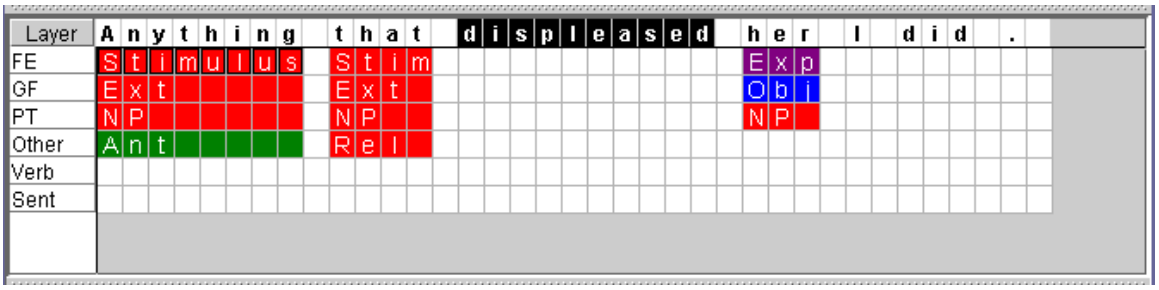

Figure 3.4: Annotation of a target in a *that*-relative clause

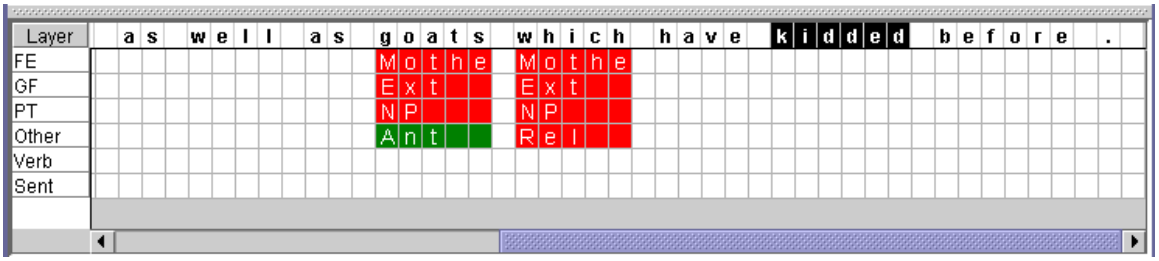

Figure 3.5: Annotation of a target in a which-relative clause

Note that, in contrast to the examples above, the relative word is not always by itself a phrasal constituent. As Figure 3.6 shows, the constituent containing the relative word may be complex.

Similarly, the antecedent phrase may be a rather complex phrase, as shown in Figure 3.7.

When a target occurs in a relative clause without an overt relativizer, as in Figure 3.8, we only annotate the antecedent phrase and mark it with the label ANT on the Other layer. Since we do not assume any kind of zero or non-overt relativizer, the label REL is not applied anywhere on the Other layer.

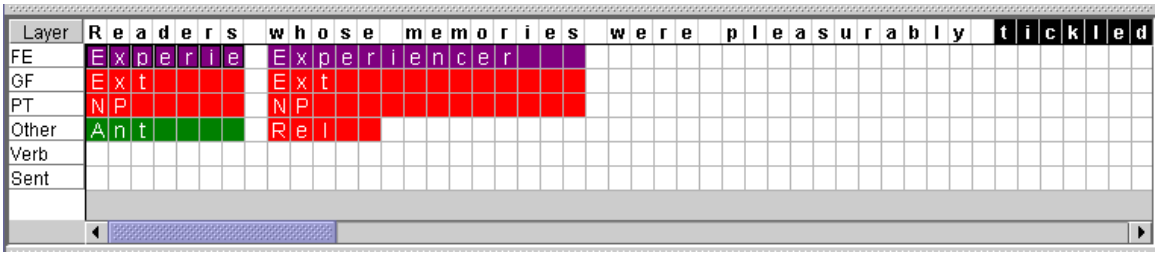

Figure 3.6: A relative phrase containing 'whose'

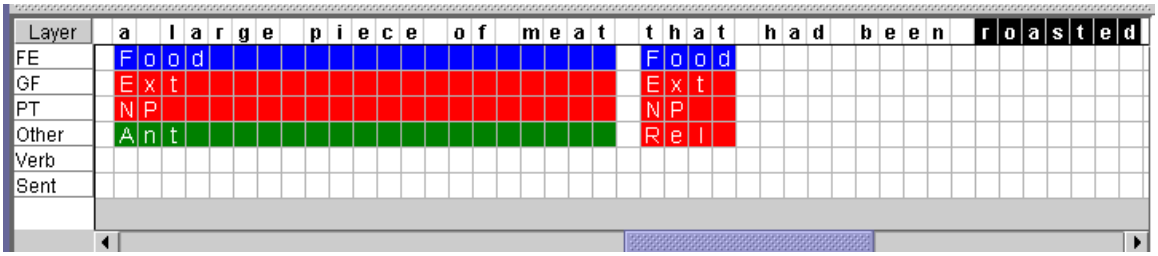

Figure 3.7: A relative clause with a complex antecedent phrase

If a relative phrase is governed by a preposition, we end up with identical FE/PT/GF triples applied to antecedent phrase and the prepositional phrase containing the relative word; the ANT and REL labels are applied as usual, as shown in Figure 3.9. In sentences with preposition stranding, the same FE/PT/GF triple occurs on the antecedent phrase, the noun phrase containing the relative word (if there is one), and the preposition. The preposition never gets a REL label assigned to it even in cases where there is no overt relativizer, as in Figure 3.10.

Notice also that the above principles for relative clauses carry over to Gov-X annotation. When the target noun is the antecedent for a relative phrase that is an argument of a verb annotated as a governor, we split the antecedent and relative phrases in the same ways as illustrated for cases of normal frame element annotation and also apply ANT and REL labels on the Other layer in the usual fashion. An example of a governor occurring in a zero-marked object-relative clause modifying the target noun is given in Figure 3.11.

Finally, infinitival relative clauses with an overt relativizer are treated just like finite relative clauses with an overt relativizer, which is illustrated in Figure 3.12.

Infinitival relative clauses without relativizer, by contrast, receive no ANT-REL marking at all. For instance, in *books to* ***read*** *over the break* the head that the relative clause modifies does not receive an ANT label.

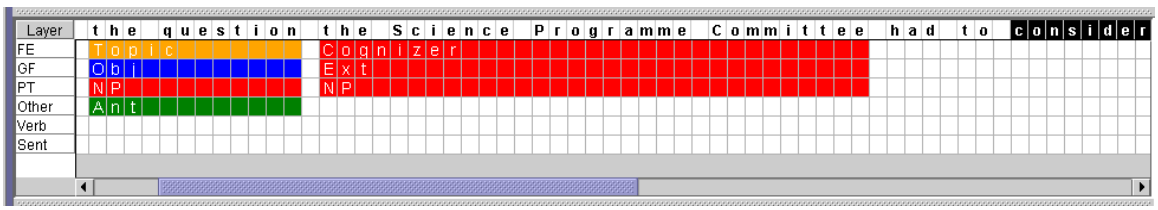

Figure 3.8: A relative clause without relativizer

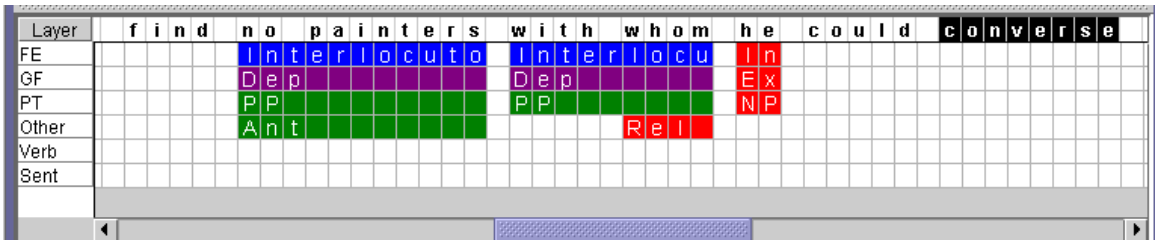

Figure 3.9: A relative clause with a relativizer governed by a preposition

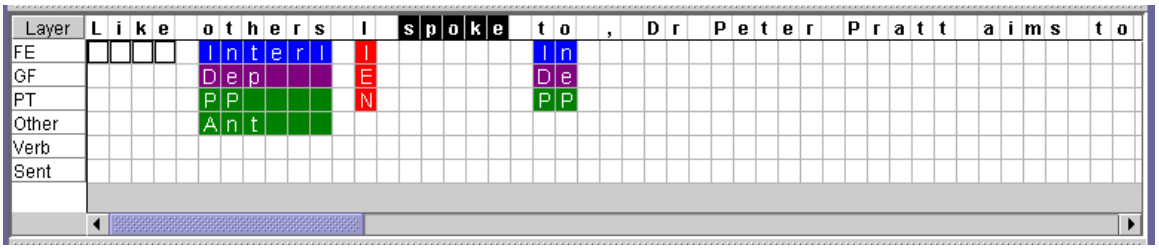

Figure 3.10: A relative clause with preposition stranding

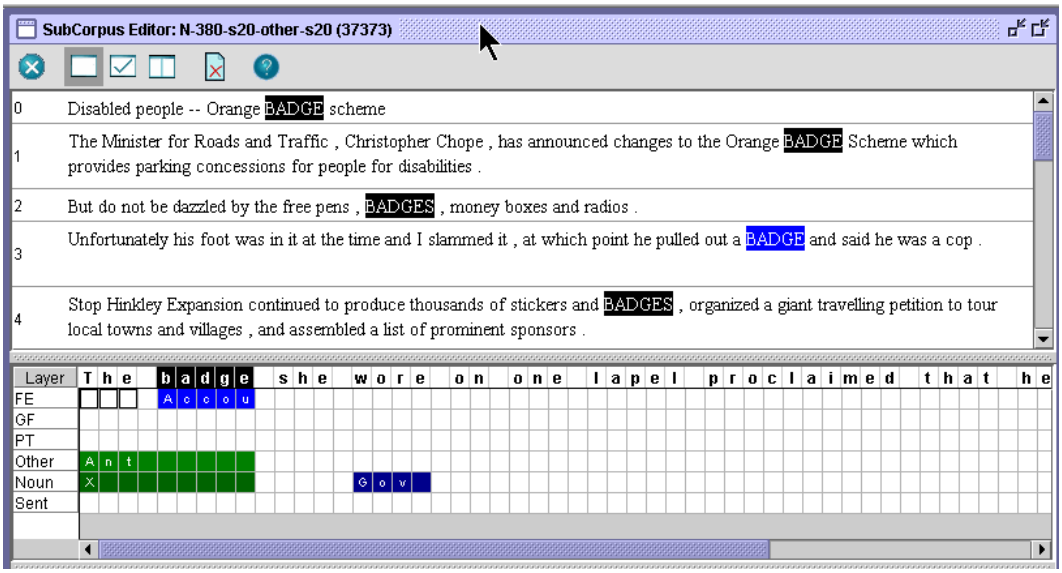

Figure 3.11: Gov-X annotation with the governor inside a relative clause modifying the target

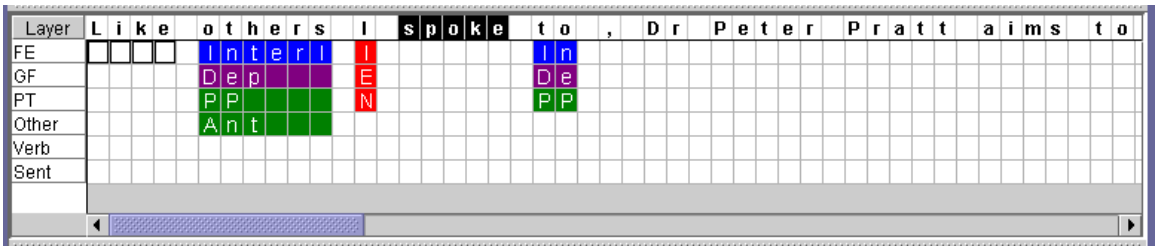

Figure 3.12: An infinitival relative clause with overt relativizer

3.2.7 Governing verbs of target nouns, adjectives, and prepositions

When annotating verbal targets, we do not record any predicates that may govern them. However, in the case of nominal, adjectival, and prepositional targets there exist several special classes of syntactic governors that we want to keep track of for lexicographic reasons.

In the case of SUPPORT expressions, COPULAS and CONTROLLERS, it would have been theoretically justifiable to omit selecting phrases outside of the standard **subcategorization frame** of the target noun and to instead rely on automatic tools for syntactic analysis to identify phrases outside the target's maximal projection that give information about the filler of a frame element role. However, since one of our goals is to provide a database that includes samples of phrases capable of satisfying particular FE requirements of the analyzed words, our decision was to increase the scope of our annotation instead. A welcome by-product of this decision is that the FN database can also serve as a resource for identifying the Support verbs and prepositions, COPULAS, CONTROLLERS, and X-Governors that FN annotators often find accompanying particular noun, adjective, and preposition targets.

3.2.7.1 Support predicates

As noted at the beginning of section (3.2), we have a special treatment for sentences in which the syntactic and the semantic head of a clause are different and where a noun target is the semantic head of the clause rather than the verb that governs it syntactically. In these cases, one or more syntactic core arguments of the support verb are necessarily understood as participants in the event or relation evoked by the target noun. These verbal arguments–typically the subject, in some cases the object, in others both the object and the subject–are tagged with labels appropriate to the noun's frame. Examples (54)-(56) exemplify support verb constructions:

(54) Aloha Gang, [Someone SPEAKER] `made` a **statement** about my need two kerrect my shpelling in ze last newsletter.
(55) [Frances Patterson PATIENT] `underwent` an **operation** at RMH today and is expected to be hospitalized for a week ore more.
(56) [One of them] `became` my **successor** in the professorship in the University of Michigan and the presidency of Cornell.

By contrast, the verb-noun combinations in (57)-(59) do not involve support structures.

(57) Did you read about his latest **mishap** in the newspaper?
(58) A senior nurse observed the **operation**.
(59) John congratulated the new **president**.

In all three examples (57)-(59), the verbal predicate governing the target noun introduces a distinct event: reading about a mishap is completely independent from participating in it (57); observing something is independent from participating in it (58); and congratulating somebody is independent from the achievement at issue (59). (Actually, relational nouns like *president* never take support verbs, they project clauses only in combination with copular verbs such as *be* or *become*.)

Support verb+noun constructions are not to be equated with idioms. While support verb-noun combinations, too, may involve some measure of non-compositionality, it is normally much less than with true idioms whose meanings cannot be built up straightforwardly from the normal meanings of their parts (e.g. *give walking papers/a pink slip/the boot*). To classify a verb-noun combination as a support verb+noun construction, it is necessary for

- the noun to denote a state, event, or relation *by itself*,
- the verb to syntactically govern the noun,
- the verb not to have the same meaning independently of the frame evoking element (more or less equivalently, this can be thought of as the frame evoking noun selecting the support),
- the verb to have very little meaning on its own, i.e. the meaning of the combination of support and noun should come virtually all from the noun

Note that elements tagged as Support ultimately will also receive a separate treatment in their own right. Thus, any differences that exist between support predicates (including most saliently the introduction of causation) will be captured by describing these predicates as frame evoking elements in very generic frames and adding the semantic type Support (see further under example 13) to indicate their limited use. In this spirit, the verb lift as used in the sequence 'lift the UN sanctions on the country' was included in the Cause_to_end frame alongside the very general LU end.v.

Verbs are not the only part of speech that can 'support' a noun. In some cases, prepositions combine with nouns to yield phrases that behave like predicative adjectives. That is, they can post-modify a head noun, as in (60), or combine with a copular verb to yield finite verb phrases, as is shown in (61)-(63).

(60) The "possessor" is the person `in` **possession** of the premises.

(61) Are health care workers `at` **risk** of getting HIV on the job?

(62) Soon, I was `in` **possession** of two dozen Eagles cupcakes decorated with white icing, green sprinkles and little plastic footballs and Eagles helmets.

(63) Some people might think that's `out of` **line** with our "democratization" policy.

3.2.7.2 Copular verbs

We may think of *be* and a few other verbs such as *appear*, *seem*, *look* etc. as a special subtype of support verb with a very minimal semantics when occurring in constructions of the form:

(i) NP_1 **Verb** NP_2/AJP/PP

Examples include:

(64) John **is** a sailor.
(65) This **seemed** a rather redundant effort to many.
(66) Tom **appears** smart enough.
(67) Massu **looked** without energy, he looked defeated seated with the towel on his face.
(68) Smithers **is** the vice-president of the armchair division.
(69) Sue **is** the mayor.

Appearing in sentence frame (i), these verbs are traditionally called *copulas* or *linking verbs*. *Be* occurs in many other sentence frames too and there are some cases of structural ambiguity where it is a copular verb under one reading, and an auxiliary under the other. In *His pastime is annoying the girls*, one can understand *be* either as a copula that pairs the role noun with the role filler, or as an auxiliary of the verb *annoy* used in the present progressive.

Note some uses of *be* with a target noun do not bear the label Copula. For example, when *be* occurs as part of the existential construction, it is tagged with the Exist label. In future data releases, *there be* occurring in the existential construction will be treated as multi-word Supp (cf. section 3.4.1.2).

As shown below in Figure 3.13, the label Copula appears on the part-of-speech specific layer of a (non-verb) target, in this case the adjective *old*.

3.2.7.3 Controllers

Recently, FrameNet has begun to recognize a new category of syntactic governor called **Controller**, abbreviated in the data as Ctrlr. It covers verbs like *merit* and *offer* when they govern event noun targets, as in (70) and (71), as well as verbs like *consider* and *find* when they govern adjectival targets, as in (72) and (73). While these predicates introduce a distinct event from that of the target, they do share a frame element with the event of the target. For Controllers of noun targets the shared frame element is typically the subject of the Controller; for Controllers of adjectives of the shared frame element is typically the object of the Controller. The constituent expressing that shared participant is labeled with a frame element relative to the noun target.

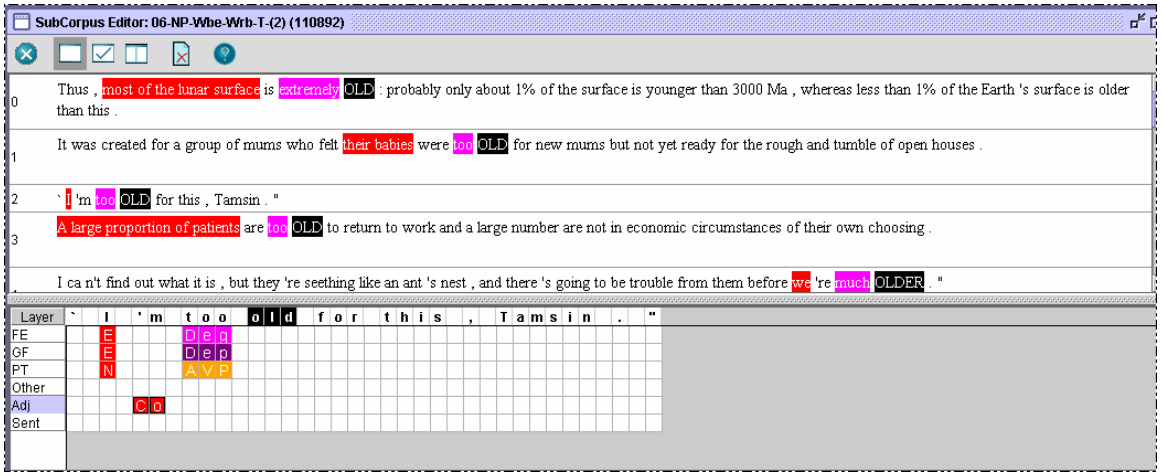

Figure 3.13: Annotation of Copula for Target *old*.a

(70) [What I paid and the satisfaction received] `merits` high **praise**.

(71) [The minister] `offered` **help** to get the various agencies coordinated.

(72) She understood that he `considered` [the trip] too **expensive** for them both.

(73) I `found` [her scenes] very **funny**.

In (70) the complex NP *What I paid and the satisfaction received* is tagged as the REASON frame element of the Judgment_communication frame evoked by *praise* and in (71), *the minister* is tagged as the HELPER frame element of the Assistance frame evoked by *help*. In (72), the NP *the trip* is labeled as the FE GOODS in the Expensiveness frame evoked by *expensive*. In (73), the object of the CONTROLLER *find*, the phrase *her scenes*, is labeled as the STIMULUS of the Subject_stimulus frame evoked by *funny*.

The CONTROLLER label is applied on the Noun and Adj layers just like the SUPP label is.

3.2.7.4 X-Governors

In the context of slot-filler annotation (see 3.8), FrameNet also uses a category **Governor**, which marks predicates that have a semantic connection to the qualia structure of the target *artifact* nouns they govern. For instance, *stab* is treated as a Governor of *knife* in the Weapon frame. The dependent constituent of the Governor that is headed by the target noun is called X for lack of a better term. No dependents of the Governor other than the X constituent are annotated relative to the Governor. Any frame elements of the frame evoked by the artifact noun that are realized within the X-phrase are annotated with frame elements as usual. As with other kinds of special governors, the Governor label, as well as the X label, are applied on the part of speech specific Noun layer.

3.2.8 Discontinuous frame elements

In some cases, the same FE label appears multiple times relative to a given target. There are two cases: multiple separate instances of the same frame element, as when several Path segments are described for a motion event (*Josh **ran** [across the meadow PATH], [along the creek PATH] and right up to Bill's barn*); a single instance of a frame element is realized in two discontinuous pieces, rather than as a single constituent. Here, we consider the latter type of discontinuous FE.

- **Raising predicates** (e.g. *may* or *might*) specify that their sole semantic argument be realized in two pieces, one as the surface subject of the clause headed by *might* and the other as a verb phrase complement of some kind. The verb phrase may be of the type bare verb phrase (VPbrst) as with *might* in (74) or it may be *to*-marked (VPto), as with *seem* in (75).

  (74) [The reader NP.Ext] **might** [want to excoriate the paper for its coverage of the so-called Clear Skies policy VPbrst.Dep].

  (75) [Old Europe NP.Ext] **seems** [to drag her feet VPto.Dep].

The FrameNet treatment does not assume any syntactic derivation of such structures from underlying forms where there is only a single, contiguous frame element. It is a fact about the semantic representation of *might* that its subject must be understood as an argument of the predicate expressed by its VP complement, rather than a syntactic fact. Accordingly, both pieces are treated as separate syntactic arguments of the raising verb and marked separately with GF and PT values.

- The **prepositional passive** construction shown below also produces discontinuities.

  (76) [The Peacock Throne NP.Ext] has been **sat** [on PP.Dep] by the Iran monarchy since the days of Nadir Shah.

  This construction belongs to the set of subject-selection constructions, which also includes the English middle construction (*This car **drives** well*). Since the prepositional passive construction is specifically geared towards selecting a non-core frame element as subject, we assign the fronted subject NP the grammatical function Ext and the phrase type NP, while the remaining preposition is tagged like a normal prepositional phrase with grammatical function Dep and phrase type PP.

- Another class of discontinuities results from various kinds of fronting constructions which sometimes lead to **preposition stranding**.

  (77) [Where she had been PP.Dep] she didn't **talk** [about PP.Dep] (Topicalization)
  (78) [The man PP.Dep] [that PP.Dep] I **lent** my phone book [to PP.Dep] made off with it. (Relativization)
  (79) [What PP.Dep] was Bill **angry** [about PP.Dep] when he got the back spasm? (Wh-extraction)

  Unlike the prepositional passive, the constructions above do not place the extracted NP into a core syntactic slot and so both pieces are assigned the same GF. We have also decided to apply the prepositional phrase type label to the displaced phrase, which is usually a noun phrase but not always, as is shown by the topicalization example in (77).

- Discontinuity is also produced by what we may call **heavy modifier shift**. This involves finite relative clauses that are separated from the nominal heads they modify and displaced appositives, as shown in (80) and (81).

  (80) About this time [a guy NP.Ext] **appeared** [who had only a pair of boxers on NP.Ext].
  (81) [Mrs Fiorentina NP.Ext] **spoke** next, [the chairwoman of Huelit-Peccard NP.Ext].

  Both pieces get the same grammatical function and phrase type values.

  Note that we do not treat the following examples as heavy modifier shift; given the assumed verbal targets, the adjectival and prepositional phrases included in curly braces {} would simply be labeled Depictive.

  (82) [Mr Burns NP.Ext] **stepped** out of the shower, {buck-naked}.
  (83) [Smithers NP.Ext] **showed up** later {with a funny hat on}

  The decision to proceed in this way is motivated by two considerations. First, Depictives express a temporary state of a participant during an event; no such semantic restriction applies to displaced appositives or relative clauses. Second, Depictives are generally more flexible with regard to their position than displaced appositives and relative clauses, as shown in the following examples:

  (84) [Looking stunned Depictive], Mr Burns **dialed** his secretary.
  (85) Mr Burns **dialed** his secretary, [looking stunned Depictive].
  (86) Mr Burns, [looking stunned Depictive], **dialed** his secretary.
  (87) *Who looked confused, Mr Smith **came** in.
  (88) *The former mayor of Denver, Wellington Webb was **spotted** at the buffet.

  For more discussion on Depictives, please consult Chapter A .

- The last class of discontinuity consists of modifying structures in which, in addition to the modified head noun or verb, the modifier takes a second argument that is realized following the head. At least with some modifiers such as *similar* in (90), the second argument is subcategorized by the pre-head modifier (rather than licensed by a construction). [9]

  In these cases, both pieces of the modifying structure fill the same grammatical function relative to the head and we assign them the same GF value, Dep. However, we treat the two pieces separately in terms of their phrase type, as is shown by the following examples. (The relevant words and phrases are annotated with triples of FE, grammatical function and phrase type).

  (89) Quite bluntly, you're a [dumber DESCRIPTOR.Dep.AJP] **man** [than I DESCRIPTOR.Dep.PP]!

  (90) The general has committed a [similar DESCRIPTOR.Dep.AJP] **blunder** [to our entry into the eastern valley DESCRIPTOR.Dep.PP].

  (91) Yet what your school did to them [so DEGREE.Dep.AVP] **irked** them [that they'd rather pay the new school than pay the old school DEGREE.Dep.Sfin].

  (92) Real people sometimes prove to be [more DEGREE.Dep.AVP] **annoying** [than anyone anticipated DEGREE.Dep.Sub].

  We assign separate phrase types to each piece here for a purely technical reason: the chunk parser that is used as part of the process that extracts example sentences from the full BNC corpus does not produce the correct result if the two pieces share the same phrase type.

3.3 Annotation with verbs as targets

Frames can be evoked by words in any of the major lexical categories of noun, verb, and adjective, as well as by adverbs and prepositions. We will begin our discussion with verbs.

3.3.1 Easy cases

Annotation is easiest when all and only the **core** frame elements (the conceptually necessary participants of the frame that a syntactic governor evokes) find syntactic expression in the sentence as separate immediate syntactic dependents of the governor. Under such circumstances, we simply annotate each of the syntactic dependents for the three kinds of information: Frame Element (that is, semantic role), Grammatical Function, and Phrase Type.

The different kinds of information are recorded on separate **annotation layers**. The separation of layers makes it possible to represent many complex situations, such as when the constituent that realizes one frame element is contained within the constituent that realizes another, or when the semantic and syntactic constituency don't match. Usually, however, most tagged constituents consist of triples of information, with coterminous tags on three layers, and they are displayed as such in the FrameNet Desktop. Figure 3.14 is a screen shot of the middle window of the annotation software in which an example sentence has been annotated. The names of the layers appear in the left-most column of the bottom frame: **FE** (Frame Element); **GF** (Grammatical Function); and **PT** (Phrase Type); **Other** (labels that deal with a small set of special syntactic constructions); and **Verb** (a layer named after, and with labels specific to, the part-of-speech of the target)

In practice, annotators only need to apply a Frame Element label; Grammatical Function and Phrase Type are derived algorithmically based on position relative to the verb and patterns of part-of-speech labels, but may require manual correction.

In addition to core frame elements, we also provide annotation for some **non-core** frame elements expressed in the sentences selected for annotation. Non-core frame elements are not necessary conceptually in the sense that they do not uniquely characterize the frame. Thus, in most frames denoting events or

[9] A modifying structure which involves subcategorization of one piece by the other cannot always be realized discontinuously. Whether or not this is possible is a lexical property of the subcategorizing predicate. Compare *similar*, as used in (90), with *eager* and *interested*.

(i) *Ronald Frump is an [eager] **man** [to affix his name to every conceivable monument]

(ii) *Please forward this email to [interested] **people** [in chess].

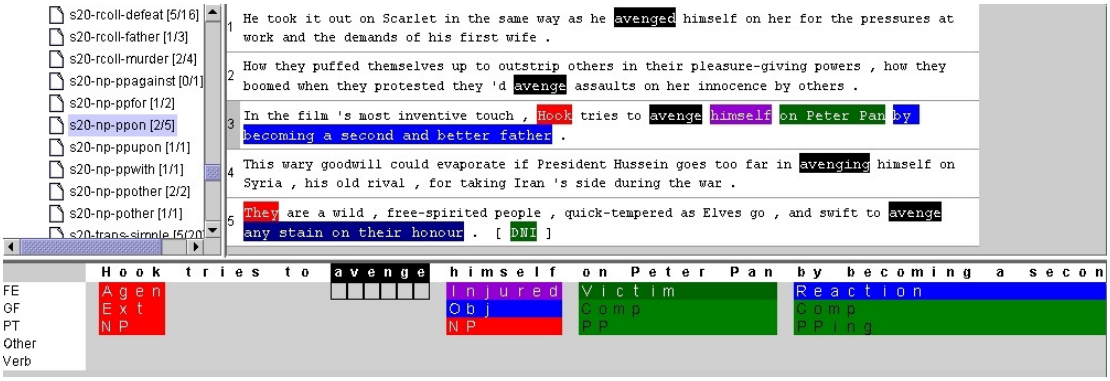

Figure 3.14: Annotation window with target LU *avenge*.v

processes, TIME and PLACE frame elements are not core, and therefore may not always be annotated. (For a more complete discussion of core vs. non-core, see Section 3.2.1). Similarly, actions often have a PURPOSE that the AGENT intends to accomplish by performing the action indicated by the target word. However, having a purpose doesn't usually distinguish types of actions and so PURPOSE is often a non-core frame element, as in the following example:

(93) They **wake** you up [to give you a sleeping pill PURPOSE].

Grammatically, non-core elements cannot be nuclear arguments (subject or object) of target verbs, and they frequently are adverbs or prepositional phrases.[10]

Unfortunately, not all sentences are as straightforward to annotate as the ones we have seen so far. We will now discuss how FrameNet deals with various challenging cases.

3.3.2 Expletives

Some syntactic constructions require the presence of non-referential material in an argument position, even though the non-referential material has no semantic relationship to the predicate.[11] In a subset of cases, this happens while a semantic argument is dislocated to a non-canonical position. The non-referential items appearing in such constructions are called **expletives**, with *it* and *there* being the two kinds of non-referential NPs in English. They never instantiate frame elements, so are not not given GF and PT tags. We do, however, record their presence with the **Null** label on the Other layer. Some typical instances of expletives are given in the following examples.

Subject extraposition

(94) [It NULL] is **clear** that we won't finish on time.

Object extraposition

(95) I **hate** [it NULL] when you do that.

Existential construction

(96) [There NULL] are more **hats** in the closet.

Subject requirement of zero-place predicates

(97) [It NULL]'s **raining**.

[10] An exception to this regularity are certain earlier noted subject selection constructions that, for example, license sentences such as *That key opened the door right away*, where the subject of the target *open* is an Instrument, typically a peripheral FE.

[11] We recognize that some traditions of linguistic analysis, such as Cognitive Grammar, do not share the judgment of non-referentiality.

3.3.3 Aspect

When verb targets co-occur with particles that are used productively to indicate aspectual information, we tag the particle with the label **Aspect** on the 'Verb'-Layer. Some examples follow.

(98) They were **chattering** [away ASPECT] in the kitchen, when the door bell rang.

(99) Mo **talked** [on and on ASPECT].

However, particles that express aspectual meanings and simultaneously form entrenched lexical units with the verb are not treated in this way. For instance, in examples such as (100), *pull through* would be treated as one lexical unit, rather than as an instance of the verb *pull* accompanied by aspectual *through* since the verb cannot occur by itself with the appropriate meaning.

(100) It's amazing what they can do these days and luckily the sick dog **pulled through**.

3.4 Annotation with nouns as targets

FrameNet also annotates relative to noun targets. In the course of our work, we have analyzed several types of nouns: nouns that denote events such as *withdrawal* and *replacement*; relational nouns such as *brother* or *girlfriend*; artifact nouns such as *house* and *vest*; and some others. Of these, event nouns and relational nouns are most clearly frame-evoking. Many of the issues discussed earlier for verbs, e.g. conflated frame elements, discontinuous frame elements, or coreness statuses apply to nouns in the same or in similar ways as to verbs. However, the annotation of nouns also brings some new challenges to the annotation enterprise.

3.4.1 Special governors

One of the major challenges in annotating nouns is that many of them participate in lexicographically relevant relationships with their governing predicates. These relationships are of four major types. Support predicates are governors of event-denoting nouns that serve mainly to project a clause centered on the frame of the noun. Existential predicates occur with event noun targets in the *there*-existential construction. Copular verbs are a semantically very bleached type of support expression used to project clauses centered on the frames of event or relation-denoting nouns. Controllers refer to an event that is separate from the one denoted by the target noun but which shares a participant with it. X-Governors of artifact nouns are verbs that evoke frames that involve the qualia structure of the artifact nouns.

3.4.1.1 Support expressions

We define support verbs as those verbs that combine with a state noun or an event noun to create a verbal predicate, allowing arguments of the verb to fill the slots of the frame elements of the frame evoked by the noun. Consider some examples from the Revenge frame.

(101) The Americans must have felt as if he was [taking SUPP] **revenge** on them for what had happened.

(102) King Menephta [took SUPP] awful **revenge** on a Libyan army he defeated around 1300 BC

Both of these examples report an act of revenge rather than an act of taking, the frame evoked by the noun *revenge* clearly being dominant. We treat verbs like *take* in a special way by marking them as support verbs on the Noun layer. The intuition behind this treatment is basically that support verbs do not introduce any significant semantics of their own. Constructions of nouns with their support verbs denote the same state of affairs that would be denoted by the noun occurring by itself.

In the examples above, recognizing *take* as a Support verb allows us to annotate its subject as the AVENGER frame element. However, a frame element of the target noun that a support verb expresses does not have to be realized as the subject of the Support verb. Objects and other complements of support verbs may also express frame elements of the target word. This last fact allows us to circumvent the problem of certain disputed constituency decisions that come up in the case of support-verb constructions. For instance, in example (103) it is contestable whether the phrase *to the press* is a complement of the noun *statement* or the verb *write*. (In contrast, with the support verb *make*, the issue would not arise as *make* by itself does not take *to*-PP complements.)

(103) He [wrote SUPP] a **statement** to the press about the bribery case.

Regardless of the analysis, FrameNet tags the phrase *to the press* with the FE ADDRESSEE as a complement of the target word *statement*. By allowing this phrase under either of the disputed analyses, we avoid the constituency decision completely.

The practice of annotating support verbs not only allows us to annotate their subjects as frame elements. It is also lexicographically necessary to record them. One reason for this is that support verbs are selected by the noun, rather than the other way around. For instance, while the noun *question* can take the support verb *pose*, other nouns in the Questioning frame such as *query* and *inquiry* take the support verbs *make*. Support verbs also vary with the sense of the noun, that is, a noun may take different support verbs depending on the frame it belongs to. Consider that in (104), the noun *argument* takes the support verb *have*, and has a meaning related to conversation, whereas in (105), the noun takes the support verb *make*, and has a meaning related to reasoning.

(104) John and I [had SUPP] a terrible **argument** last night.

(105) John [made SUPP] a convincing **argument** that the project should be funded.

With regards to their semantic contribution, we recognize several types of support predicates:

- Plain Vanilla: the support adds virtually nothing to the frame evoking element (e.g. *make a statement*)
- Aspectual: the support changes the temporal focus of the event portrayed by the frame evoking noun, e.g. *start* in *start an operation*; this also covers things like *get/go/fall into a (foul) mood* vs. the vanilla support structure *be in a (foul) mood*.
- Point-of-view: the support changes the profiled point-of-view of the frame evoking noun, e.g. *undergo* in *undergo a physical exam* (the patient's point of view) vs. *give a physical exam* (the doctor's point of view)
- Registral: the different support verbs appeal to different formal registers, e.g. *make a complaint* versus *register a complaint*; *take revenge* versus {*exact/wreak*} {*revenge/vengeance*}.
- Causative: the support adds another participant and the idea of causation to the basic scene. These generally occur paired with a non-causative support, as in *put in a (foul) mood* versus *get into a (foul) mood*; '*bring into play* vs. *come into play*'; *give a headache* versus *have a headache*, etc. Note that normally we only tag the causee, e.g., the object of *give* or *put*, as a frame element of the frame evoked by the target. Additionally, we tag the subject of the support verb when it fills a frame element role that is also part of the basic frame. This is the case, for instance, for the COMMUNICATOR frame element occurring with the noun *call* in the Contacting frame in sentences such as *I 'll give you a call when Mr Steen is back in town.*[12]

Note that the above classification is an informal one that is not encoded in the database through the use of different support labels. Further, the above division is clearly much less fine-grained than the distinctions that exist among **Lexical Functions** in the sense of Igor Mel'cukMel'čuk (1996). An extension of FrameNet could be defined which sought to distinguish the full range of Lexical Functions.

In addition to support *verbs*, we recognize a second type of support expression, namely combinations of nouns with support *prepositions*. Support prepositions combine with certain target nouns to yield a phrase that is more or less equivalent to a predicative adjective.[13] As in the case of support verbs, the frame of the noun is dominant and it is the noun that selects the particular support preposition, rather than the other way around. For instance, while the noun *danger* in the Run_risk frame is supported by the preposition *in*, the noun *risk* in the same frame requires the support preposition *at*.

[12]The treatment of causative frame elements is not quite satisfactory in that with certain support verb+noun combinations, a frame element that is peripheral ends up occupying a core syntactic slot, as is the case for *us* in (i) and for *Carl* in (ii), respectively.

(i) [Give SUPP] [us] a **grin**, kiddo!
(ii) [Carl] [gave SUPP] Linda the **nickname** Butsy.

We also would like to be able to mark arguments added by the support verb, especially when the composite frame resulting from the combination of verb and noun is a well-defined frame we have in the FrameNet database.

[13]We might want to recognize another possibility: a preposition may turn a noun into an adverb, e.g. *at X's risk/peril*

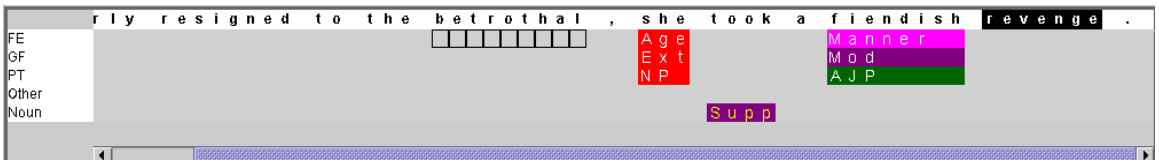

Figure 3.15: Annotation of Support Verb for Target *revenge*.n

When a target noun occurs with a support preposition and projects a finite clause, we also apply the Copula label to the copular verb that combines with the preposition and noun to form a verb phrase. Some examples are given below.

(106) The painting [is COP] [on SUPP] **loan** from Mr Smithers

(107) About 650 oil wells [are COP] [on SUPP] **fire**.

(108) This old book [is COP] now [in SUPP] my **possession**.

(109) The documents [came COP] [into SUPP] my **possession** by inheritance.

Notice, though, that there need not be a copular verb for a target noun to occur with a support preposition. When the combination of support preposition and noun is used as a depictive or resultative secondary predicate, it is not accompanied by a copular verb. Illustrative examples from the Facial_expression frame are given here.

(110) His face was twisted [in SUPP] a **grimace**.

(111) He wrinkled his brow [into SUPP] a **frown**.

Usually, such sentences are roughly parallel to sentences with a support verb. For the two preceding examples, *make* is a good candidate.

(112) His face was twisted, [making SUPP] a **grimace**.

(113) He wrinkled his brow, [making SUPP] a **frown**.

Finally, notice that in accordance with their status as syntactic governors for target nouns, support expressions are not tagged with frame element labels. They are tagged on the Noun layer of the target nouns and have no grammatical function or phrase type assigned to them, as shown in Figure 3.15.

3.4.1.2 Existential sentences

A subset of nouns can occur in a *there*-construction that serves to introduce an instance of the frame evoked by the noun target into the discourse. FrameNet so far has recorded the governing verbs occurring in the *there*-construction by applying the label **Existential** (abbreviated EXIST in the data) to the relevant verb on the Other layer. Usually the verb is *be*, but others, such as *come*, *occur*, and *ensue* are found as well. Examples are given in (114) and (115).

(114) There [was EXIST] an **argument** between the players.

(115) There [ensued EXIST] a **discussion** about the current program.

The word *there* that occurs before the existential verb is marked with the NULL tag, which like the EXIST tag is applied on the Other Layer, as is shown in Figure 3.16 .(See section 3.3.2 for more information on the NULL tag).

The EXIST label was introduced earlier and separately from the SUPP label. In the release following Release 1.3, we intend to convert all instances of the EXIST label to SUPP. The reason for doing so is that with the exception of *be* all other verbs occurring in the *there*-existential construction can also occur as SUPPORT verbs of nouns outside of the *there*-construction, as is shown by examples (116) and (117).

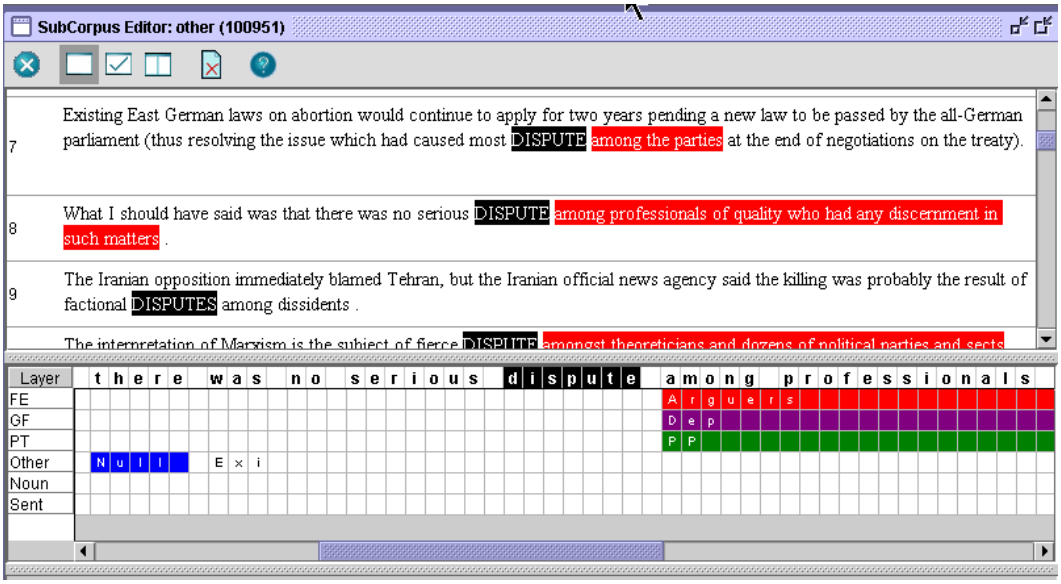

Figure 3.16: Use of Existential label with event noun target

(116) Another private wrestled the revolver away from Czolgosz while a **melee** ensued.

(117) The **robbery** occurred on Smithfield road six miles east of Campbellton.

The combination of *there* and *be* will be treated as an idiomatic multi-word support–it is actually the same as the lexical unit *there be* in the Existence frame. That is, in *there be*, the word *there* will in the future be labeled as Supp on the Noun layer rather than as Null on the Other layer.

3.4.1.3 Copulas

As pointed out above in section 3.2.7.2, verbs like *be*, *seem*, *appear* and a small number of others can be considered to be a special, rather bleached kind of support verb. However, unlike the semantically heavier support verbs, copular verbs can occur with relational nouns such as *chairperson* (118), not just with event-denoting nouns such as *proposal* in (119).

(118) Jeanne Hossenlopp will be the next **chairperson** of the department.

(119) His **proposal** was that Carthage should be destroyed.

Copular verbs occur in two distinct constructional contexts. This is illustrated by the examples in (120)-(123).

(120) Jim **is** a tailor.
(121) This **seemed** a silly idea to me.
(122) Jeidels **is** the director of the physics department.
(123) Paul **is** the treasurer.

The four example sentences, though superficially the same, instantiate two different constructions. The first two examples, (120) to (121), are predicative or ascriptive. In (120), the property of being a tailor is ascribed to Jim; in (121), the property of being a silly idea is ascribed to the action anaphorically referred to by *this*. The second use, illustrated in (122) and (123), is specifying or equational. In (122), the identity of the referents of *Jeidels* and *the director of the physics department* is asserted; in (123), the identity of Paul and the filler of the treasurer role is asserted. In specifying sentences, NP_1 and NP_2 can be switched, with the resulting sentence still meaningful and even having the same truth conditions. This is not possible with predicational sentences, as is shown by (120') and (121').

(122') The director of the physics department **is** Jeidels.
(123') The treasurer **is** Paul.
(120') *A tailor **is** Jim.

(121') *A silly idea **seemed** this to me.

FrameNet annotation practice does not distinguish the above two constructions.[14] *Be* is tagged as copula in both cases.

3.4.1.4 Controllers

Controllers are the third major kind of special governor that FrameNet records. They are different from Support expressions and Copular verbs in that they evoke a separate frame from the one evoked by the target noun while still having a shared participant. Consider examples (124) and (125).

(124) In this case it means that the bad deed deserves **revenge**.

(125) Bill offered **help** in the case of an emergency.

Example (124) is not a report of actual or imagined revenge. It is only asserted that revenge is warranted for the bad deed referred to. However, the *bad deed* clearly would be the INJURY that an act of revenge would be intended to punish. The situation is similar in (125): offering help is not the same as actually helping. But somebody who offers help would have to deliver help, that is, be a HELPER in an actual instance of Assistance, if called upon.

When a **Controller** verbs like *deserve* and *offer* is present, then the constituent representing the shared participant, typically the subject of the Controller, is annotated with a frame element label relative to the noun target. Just like the Supp label the Controller label is applied on the (part of speech-specific) Noun layer.

(126) The Prime Minister encouraged **revenge** against Absurdistan.

(127) We welcome the Elbonian **decision** to withdraw behind the lines of 1328.

Example (126) also does not assert an instance of the event evoked by the target noun. Encouraging somebody to do something is different from doing it and, crucially, the encourager need not even be a participant at all in the encouraged act. Likewise, in example (127), welcoming a decision is unrelated to making a decision and two such events normally share no participants, since people do not normally talk about the felicity of their own decisions. Verbs like *encourage* and *welcome* will most likely continue to go unlabeled, although we are considering the possibility of introducing a category **Concomitant** for verbs that evoke a frame that is related via a background scenario to the frame evoked by the event noun. For instance, to grade an exam is not the same as to give or take one. But since evaluation is a necessary part of the overall scenario of examination, it seems desirable to record the close connection between *grade* and *exam*. Another example of a Concomitant is the verb *call* when it governs the noun *election* (in the context of British politics). The person calling the election need not vote or stand for election but clearly the calling of the election is an integral part of the electoral process.

3.4.2 Noun Compounds

Our treatment of **noun compounds** recognizes, but does not easily cover, the full range of problems that they exhibit, which any theory of the grammar of English must address. In practice, we distinguish compounds that are treated as single LUs, for which FrameNet (at least currently) gives no account of their substructure (e.g. *firing squad*, *sugar daddy*, *wine bottle*), from compounds based on frame-evoking nouns whose modifiers happen to be nouns or relational adjectives (e.g. *restoration costs*, *military supremacy*). Almost all noun compounds are **motivated** in that their components have clearly defined semantic relations to each other, whether or not the pattern they exemplify is productive. In addition, the separation between conventionalized two-part words and examples of productive processes is not always clear, though there are clear cases. Compounds like *fire risk* and *health risk* would be treated with the noun *risk*, and the modifiers would be given FE labels that follow a description of the head noun's frame. Similar treatment is given for the compounds *language pedagogy*, *animal sacrifice*, and *water flow*, where the modifier is a noun, as well as *presidential privilege*, *economic problem*, *educational crisis*, etc., where the modifier is a relational adjective.

[14] In the past, we recorded a distinction, with the label SUPP for predicative uses and the label Cop for specifying uses in which the role-denoting noun preceded the copula and the filler-denoting noun followed it. We abandoned this practice because it led to terminological confusion about the meaning of *copula* and *support*, and also resulted in inconsistent annotation.

A circular piece of white lint was pinned over his heart as an aiming point and the FIRING SQUAD , of the Scots Guards , opened fire .
On March 10 Gustav Just resigned as a Social Democrat (SPD) deputy in the Brandenburg parliament after admitting that he had participated in the execution by FIRING SQUAD of six Ukrainian Jews in July 1941 .

Figure 3.17: Compound Noun as LU - *firing squad*.n

Yet , the work 's direction is quite the opposite of that conventionally assigned to the fertility RITE .
In certain fertility RITES in Burma , a woman desirous of offspring is required not only to approach the King Cobra but to plant a kiss on its mouth .

Figure 3.18: Compound Noun as Head Noun+Modifier - *fertility rite*.n

Note that **in the context of noun-noun compounds, we never annotate the head noun as a frame element of a frame that may be evoked by the non-head**. Thus, for a compound like *weapons treaty* we annotate *weapons* as a frame element relative to the noun *treaty* (which belongs to a Documents frame), but we do not annotate *treaty* as a frame element relative to *weapons*. While the non-head must figure in some frame evoked by the head, the reverse is not true in the same way; there is no sense that *weapons* saliently and consistently causes speakers to think about how their manufacture and possession is regulated. We consider as accidental any compounds in which the head noun superficially looks as if it fills a frame element role in a frame evoked by the non-head. Thus, although in the compound *frying pan* the noun *pan* appears to express the CONTAINER frame element in the Apply_heat frame evoked by *fry(ing)*, we do not annotate the head *pan* as a frame element of the non-head *frying*. The policy exemplified here for noun-noun compounds also applies to compounds consisting of a relational adjective and a head noun. Thus, while we annotate the relational adjective relative to the head noun of the noun phrase in [*economic*] **policy** and [*military*] **might**, we do not annotate the head noun relative to the relational adjective (***economic** [*policy*]; ***military** [*might*].)

Both in the case of noun-noun compounds and of compounds consisting of relational adjective and noun, we may still annotate the head as a Governor of the non-head target (cf. section 3.8 on slot-filler annotation). Also, in a compound, we always annotate any dependents of a non-head target that occur to its left. Consider the annotation of *[food]* **processing** *facilities*, where *food* is tagged as MATERIAL with respect to the Processing_materials frame evoked by *processing* but the head noun *facilities* is unlabeled. Notice that this contrasts with our normal annotation for pre-modifying uses of non-relational, predicative adjectives. For instance, when *related* in its Cognitive_connection sense occurs as a pre-modifier of a noun, we annotate both any preceding noun that may fill its CONCEPT_1 role as well as the head that the adjective modifies: *[work]*-**related** *[stress]*.

Figure 3.17 shows our treatment of the compound noun *firing squad* as a single lexical unit. Notice that the whole compound is the target (indicated with capital letters), and is annotated with the FE EXECUTIONER. (The annotation on the target is not visible on black and white print-outs.)

Figure 3.18 shows our treatment of the compound noun *fertility rite* as a head noun modified by a noun. Notice that the target word *rite* is modified by a noun that is annotated with the FE DESIRED_STATE.

3.4.3 Null Instantiation with Noun Targets

In the case of noun targets, null instantiation is very common, and it is much more difficult to decide what licenses the absence of the noun's conceptually necessary arguments. In particular, quantification and generic use often make the notion of DNI inapplicable since they tend to prevent the individuation of particular events or states and their participants. Consider the following pair of examples, which illustrate this effect of quantification and generic construal for a verbal and a nominal target.

(128) Every time Max did something like that, I knew how to **get even**.

(129) **Revenge** is sweet.

However, even without quantification, there are no reliable clues from definiteness marking about how frame elements of a target noun are contextually known. Consider the following example sentence uttered in a courtroom context, phrased with a verbal target.

(130) Smithers was **convicted** after two hours of deliberation.

Now, consider an example with the noun *conviction*.

(131) His role in that **conviction** is now under investigation.

It is clear that in the Verdict frame, whose LUs include the noun *conviction* and the verb *convict*, the FE Charges is a core frame element. The example in (130) with a verbal target is felicitous only if the FE Charges is contextually recoverable. By contrast, a sentence like (131) with the nominal target *conviction* is felicitous even if the Charges are not fully recoverable in context. This is demonstrated by the fact that sentence (131) could be preceded by either one of the following discourses:

(131') We have reason to believe that this agent has previously tampered with evidence to get a conviction. Just two months ago the murder **conviction** of Howie Cheatham was overturned. Agent Smith also was a witness there.

(131") We know that this agent has previously tampered with evidence to get a **conviction**. There is, for instance, the case of a man in Missouri that had to be released after it was found that agent Smith had manipulated fingerprints.

In the first example, the Charges (*murder*) are explicitly mentioned. In the second, they are not identified at all, just existentially bound to the event of the earlier trial and conviction. Thus, reference to an event with a definite event noun does not require that speaker and hearer can resolve all aspects of an event to entities that they know of independently of the event talked about by the target word under consideration.

Basically, it seems to be the case that there are no frame-evoking nouns that always require the expression of some or all of their frame elements. Still, there are situations where it is clear that a particular frame element has to be interpreted as omitted under DNI. For instance, relational nouns like the kinship terms *brother*, *sister*, etc., are most commonly used to refer to specific people in non-generic contexts and when this is so, we annotate omitted relata as DNI. An example of this is given in (132).

(132) The [**brother** ALTER] was found not guilty [DNI EGO].

Likewise, when an event noun projects a finite clause in combination with a support verb and the event is not construed generically, FN records frame elements that have to be recoverable in the discourse or the discourse setting as DNI. Thus, we record DNI for the Goal frame element in example (133), which refers to a specific occasion of entering, but keep no record, not even INI, of the frame element in the habitual sentence (134).

(133) He made his **entrance** singing "Oh, What a Beautiful Mornin/"' [DNI GOAL].

(134) He made his **entrances** and his exits like the man of the stage he was and never forgot his theatrical background.

While there may be other constructional and interpretational contexts in which a non-realized frame element of a target noun clearly receives a DNI interpretation, we currently only record DNI for noun targets in the above two cases: for specifically referring relational nouns and event nouns that are used in finite clauses with support verbs to refer to specific occasions.

3.5 Annotation with adjectives as targets

Adjectives can also evoke frames, that is, be frame-evoking. Certain semantic areas such as emotion- or evaluation-related frames actually have a considerable share of adjectival lexical units. FrameNet annotates adjectives, both when they are used attributively and when they are used predicatively. For the most part, adjectives take the same kinds of phrases as dependents that verbs or nouns do: finite clauses (*Phil is **happy** [that he passed]*); infinitival clauses (*Phil is **eager** [to finish]*); gerundive forms (*Phil is **busy** [writing]*); prepositional phrases (*Phil is **curious** [about the new student]*); adverbial phrases (*Phil is [rather] **disappointed***); etc. However, predicative adjectives are different from verbs in that they do not take direct objects, with the possible exception of *worth* (as in *It's not **worth** [the trouble]*). And attributive adjectives, of course, modify nominals rather than complete noun phrases (***Eager** [customers] snapped up items as soon as they were displayed*).

As pointed out earlier, when adjectives are used predicatively, the annotation of verbs like *be* is the same as with predicate nominals: we tag them Cop(ula) on the part-of-speech specific layer, which is abbreviated Adj layer in the case of adjective targets.

3.5.1 Relational modification

There is a distinct subclass of adjectives like *economic, medical, military, judicial* that can never be used predicatively as is shown in the following examples.

(135) The White House announced a new **economic** policy.

(136) *The policy that the White House announced is **economic** .

We call these adjectives **relational** modifiers; other names used for them include **pertainyms** or **domain adjectives**. Relational adjectives are comparable to the modifying noun in noun-noun compounds (e.g. *medicine* in *medicine man*). They do not modify the referent or the extension of a head noun, but rather its sense or intension, hence the name. Dictionaries typically define them with phrases like *having to do with, relating to, pertaining to, characteristic of* [some abstract or concrete entity]. These adjectives are not frame-bearing, at least not in a way that is concrete enough to allow for a clear definition of a scenario and of a set of frame elements. Although they are placed in the frames with which they are broadly associated–the adjective *retributory*, for instance, lives in the Revenge frame–we do not provide full annotation for them, just as we do not annotate the heads of noun-noun compounds from the point of view of target nouns that pre-modify the head noun (cf. section 3.4.2).

3.6 Annotation with adverbs as targets

Adverbs also evoke frames, that is, they are frame-bearing. In many cases, adverbial evocation of a frame alternates with adjectival evocation. Compare the following pairs:

(137) Bill was **wise** to sell the piano.
(138) Bill **wisely** sold the piano.
(139) Dillon was **happy** to carry the load.
(140) Dillon **happily** carried the load.

Adverbially evoked frames are typically not the pragmatically dominant frames in a clause or sentence. For instance, whereas sentence (137) as a whole intuitively is an assessment of Bill's mental properties, (138) is foremost a report of a selling act. Support for this view comes from applying the so-called lie-test: if one challenges (137) by exclaiming *That's a lie!* one is contesting a different claim than somebody responding in the same way to (138). In the former case the wisdom of the act is challenged, in the latter case the act itself is challenged.

So far, the FrameNet project has not made adverbs such as *wisely* in (138) or *happily* in (140) targets of annotation; they were, however, annotated as dependents of the, typically verbal, semantic heads they modify.[15]

FrameNet largely concentrates on adverbs expressing speaker attitudes such as candidness or directness (cf. (141)-(142)), epistemic and evidential adverbs such as *probably* in (143), *presumably* in (144), and *reportedly* in (145).

(141) **Frankly**, my dear, I don't give a damn.

(142) **Honestly**, I find this very disturbing.

(143) In this case, Bayesian methods quickly reassure us that the Pope is **probably** not an alien.

(144) An individual **presumably** fabricated a message purporting to be from me.

(145) There are many herbal weight-loss supplements on the market today, and they utilize different mechanisms to **reportedly** aid in dieting.

Note that even if the clause that a target adverb modifies is broken up in two pieces by the adverb, the pieces still receive identical grammatical function and phrase type labels. The grammatical function will always be Head and the phrase type will mostly be Sfin. There are, however, cases such as (145), where the adverb only has scope over the infinitival purpose clause; the phrase type will then be of a different type, e.g. VPto in the case of (145).

[15] Adverbs such as *happily, carefully*, etc. may appear as frame elements of event nouns when the latter project clauses in combination with support verbs.

3.7 Annotation with prepositions as targets

Most frames do not include prepositions among the target lexical units. However, in a considerable number of frames, prepositions occur as semantically inactive markers of frame elements with some of the lexical units, as illustrated in (146) and (147).

(146) Our exit strategy **depends** on our entry strategy.

(147) Don't **Put Up** With Spam

While in each case the choice of a particular marker preposition must have some historical motivation, there is no clear understanding that the preposition evokes an identifiable frame that could also be evoked by the preposition when it heads a simple clause.

However, prepositions are appropriate targets of annotation in frames that cover the vocabulary of space, time, and motion, as shown in the following examples.

(148) The audience shouted that [Cinderella] was **in** [the cupboard].

(149) Lay [the insert] **on** [the table] and trace around it.

(150) Sue dropped [her handbag] **on** [the table].

(151) [The beer] **in** [the fridge] was no good.

We annotate sentences such as (148) and (149) relative to the prepositions *in* and *on*, respectively, which are targets in the Locative_relation frame. The phrase following the preposition denotes the LANDMARK of the relation, and it is assigned the grammatical function Obj(ect). The FIGURE of the spatial/temporal relation is tagged as an Ext(ernal) argument. Note that this so even when the prepositional phrase post-modifies another noun as in (151). The phrase types for Figure and Ground expressions are chosen as appropriate. Very often, both Figure and Landmark will be NPs.

Note that in sentences like (148), we tag the verb *be* as a COPULA on the (part-of-speech specific) Prep layer. Similar to what is found with nominal targets, some sentences with prepositions as targets allow reversal of the pre- and post-copular elements for pragmatic effect (*Inside the drawer was a smoke alarm, still unopened in its box*).[16] FN annotation does not capture these differences; we always assign the label COPULA (COP).

Prepositions also often can take modifying phrases that immediately precede them as in (152) and (153).

(152) Bob took out a brand new sleeve of balls, teed one up and put it [right] **into** the water.

(153) New Oxford is [10 miles] **before** Gettysburg.

These modifiers are assigned the GF Dep(endent) and whatever phrase type is appropriate, for instance, adverb in (152) and NP in (153).

3.8 Annotation relative to slot fillers

Some nouns–for example, natural kind and artifact nouns like *tomato*, *hammer*, or *pants*–do not evoke frames by themselves, or do so only marginally. These nouns mostly just occur as slot fillers in frames evoked by verbs, adjectives, or other nouns. Nevertheless, we tag a governing verb or preposition on the Noun-Layer as a governor for some of these slot filler nouns.

The motivation for doing this is as follows: For certain entities, we would like to know in which frames they appear as slot fillers. Conversely, we also want to know, for a particular FE, what its typical slot fillers are. For example, consider the Building frame. We might wonder what kinds of things are usually built or constructed. Starting by thinking about various kinds of buildings, we might want to know in what kinds of events they participate. We may intuit that they are often mentioned as objects of *build*, but are there other kinds of building, for instance, that apply to buildings or their sub-parts?

We could attempt to automatically derive this kind of information from our annotation relative to governors like *build*, *construct*, *assemble*, etc. However, for that to be useful, we would have to annotate many

[16] Incidentally, there are also real subject uses of prepositional phrases. The sentence *Under the desk is a good place to hide* answers the question *What is a good place to hide?*. It does not function to convey information about the presence of a type of entity (a good place to hide) in a discourse-accessible location, which is the case for the inversion example with *inside the drawer*.

more sentences for each of the governors than is current practice.[17] For lexicographic purposes, it is not necessary to document that in addition to houses, towers, and many other things, castles can be constructed or put up, and so we have not done it. Thus, currently, looking for sentences with nouns such as *castle*, *tower* etc. and treating them as targets is the the easiest way to address questions like the ones above. We can then record what syntactic governors (**Gov**) take phrases containing the artifact noun targets as arguments, and what kind of syntactic constituents contain the target. (For lack of a better term, we call the constituent containing the target slot filler noun **X**.) Examples of what we call Gov-X annotation are given below.

(154) He [built GOV] [two tall **towers** X].

(155) Design and [construct GOV] [the **castle** X].

Notice that we tag both the verb and the particle as GOV when a phrasal verb is the governor of a target noun, whether or not the verb and particle are contiguous.

(156) Leslie [put GOV] [up GOV] [the **towers** X].

(157) Leslie [put GOV] [the **towers** X] [up GOV].

In several frames containing artifact nouns, we also have defined some frame elements reflecting a kind of qualia structure of the artifacts (see Pustejovsky (1995)). For instance, we may record modifiers that denote the material from which an artifact is made (constitutive quale) or the purpose for which it is used (telic quale). Here are some examples from the Clothing frame, which contains nouns denoting items of clothing.

(158) Dot always [slept GOV] [in [her WEARER] [**vest** GARMENT] and knickers X]

(159) I [put GOV] [on GOV] [[my WEARER] [smart DESCRIPTOR] **suit** X] for the last of the interviews.

In addition to Gov-X annotation and qualia-type annotation, we can also attach a semantic type to some or all of the lexical units in a frame. For instance, we could assign the semantic type 'Dress' to all items of clothing. Similarly we could give nouns like *wood, concrete, ivory* in a *Materials frame a semantic type 'Material'. This kind of information would then give users of the FrameNet database information about classes of words that meet certain selectional restrictions.

One special kind of noun that received Gov-X annotation are what we call **transparent** nouns. Nouns like *top, pound, bunch* can appear as the first noun in N_1-of-N_2 constructions in contexts where the governing verb semantically selects N_2 rather than N_1, the syntactic head. In these contexts, we call N_1 **transparent**.

(160) Sue drank a **cup** of hot coffee.

(161) He pinned a **square** of fabric on the back.

(162) She went to see her **idiot** of a husband.

Semantically, the nouns that can be transparent fall into the following classes:

- Aggregates (*bunch, group, collection*)
- Quantities (*flood, number, scores*)
- Types (*breed, class, ilk, kind*)
- Portions and Parts (*half, segment, top, bottom*)
- Unitizers (*glass, bottle, box*)
- Evaluations (*gem, idiot*)

While these words were annotated in appropriate frames such as Aggregate in the case of *group*, *bunch*, annotators added Gov-X annotation for those sentences where the target noun appears as a transparent N_1:

(163) Leslie [ate GOV] [a [**sliver** PIECE] [of lemon SUBSTANCE]X].

(164) Robin [drank GOV] [[a COUNT] [**pint** UNIT] [of beer STUFF] X].

[17] We aim for around 3 -5 examples per subcorpus, resulting in an average of about 20-25 sentences per lexical unit.

Chapter 4

Identifying Phrase Types

As discussed in 3.2, the syntactic metalanguage used in the FrameNet project is intended as a framework for **lexical description**–i.e. to describe the syntactic valence properties of individual lexical items. It is not intended as a framework for the complete syntactic description of sentences. In choosing the phrase types and grammatical functions, the major criterion was whether or not a particular label might figure into a description of the grammatical requirements of one of the **target words**. Our goal is to annotate words or phrases in a sentence that are either in direct grammatical construction with the target word, where this notion is extended to include both extracted, extraposed constituents (cf. section 3.2.8), or that are dependents of higher commanding predicates that are construed as participants in the target word's frame (cf. section 3.2.6).

Initially, the emphasis of FN annotation was on what was most relevant to lexical descriptions, namely the core and peripheral frame elements of target words. Accordingly we limited ourselves, for the most part, to those phrase type labels needed for the annotation of such elements. Over time, annotation experience required broadening the range of elements we annotate to include extra-thematic frame elements. Such frame elements evoke frames that are distinct from the one evoked by the target, and typically embed the target frame in a larger scenario (cf. section A). While we find the inclusion of extra-thematic elements in our annotations valuable from a lexicographic point of view, and in fact necessary for the annotation of full texts, the concomitant introduction of additional phrase types has led to some inconsistencies in the number and type of distinctions made between phrase types. (These will be pointed out below). In the future we may revise the inventory of phrase types to make it more compatible with what would be desirable for theoretically-oriented syntactic descriptions.

Finally, note that there are two cases in which frame element labels are not paired with grammatical function and phrase type labels. First, although certain types of noun targets can sometimes bear frame element labels, they are never assigned phrase type or grammatical function tags, as is shown in Figure 4.1. This policy applies mainly to relational target nouns such as *brother* and *mother*, and to artifact and natural kind-denoting target nouns such as *building* and *forest*. Likewise, frame elements annotated on any other layer than the first FE layer are never assigned grammatical function and phrase type labels (cf. 3.2.5 on frame element conflation).

4.1 List of phrase types

What follows is a list of phrase types that are used in FrameNet, accompanied by some examples. The phrase types are discussed and exemplified in greater detail in sections 4.2 and higher. Phrase types are assigned automatically by the FrameNet desktop software during the annotation process, but may require manual correction.

4.1.1 Noun Phrase Types

Non-referential noun phrase

Non-referential noun phrases, also called expletives, such as *there* in *There was a row* and *it* in *It was raining*, are not assigned frame element labels and consequently cannot have phrase types either. Such constituents are marked only on the Other layer.

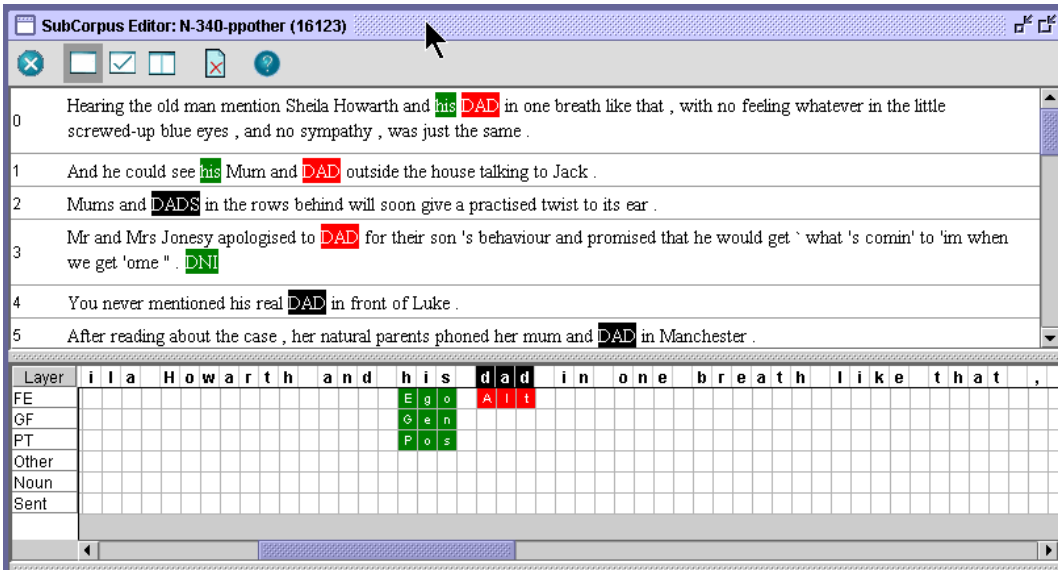

Figure 4.1: Frame element annotation of a Target word

Possessive Noun Phrase (Poss)
The type possessive noun phrase includes both the possessive determiners, shown in (1), as well as noun phrases with the *'s*-genitive clitic, shown in (2).

(1) {My, your, his, her, its, our, their} **arrival** surprised everyone.

(2) {John's, the President's, . . . } **statement** will be aired at 8 p.m.

Non-maximal Nominal (N)
Nominals that are not referentially complete, especially in compounds, are given the phrase type label N. Such nominals may themselves be internally complex (cf. (3)).

(3) [fast food] **allergy**

(4) [car] **manufacturer**

Standard Noun Phrase (NP)
Referentially complete noun phrases that could fill core verbal argument slots are assigned the label standard NP, as in the examples below.

(5) [My neighbor] is a lot **like** my father.

(6) [John] **said** so, too.

(7) [You] **want** more ice-cream?

(8) [The notebook I found] [1] to Sue.

(9) [Two women] **came** in.

4.1.2 Prepositional Phrase Types

Prepositional Phrases (PP)
Two types of Prepositional Phrases are assigned the phrase type **PP**.

- Standard Prepositional Phrase (with NP object)

  (10) **Scrape** it back [into the microwave bowl].

- Particle (with no object)

[1] belonged

(11) I carefully **peeled** the skin [off]

PPing (Preposition with gerund object)

(12) Peter **thought** [about going home]

PPinterrog
labelpt:ppinterrog (Preposition governing a *wh*-interrogative clause)

(13) He taught me not to **think** [about where I had been and what I had done].

(14) I **worry** [over why we cover this story].

4.1.3 Verb Phrase types

4.1.3.1 Finite Verb Phrase (VPfin)

(15) Who did she **believe** [had left]?

4.1.3.2 Nonfinite Verb Phrase

- **VPbrst** (Bare Stem Verb Phrase)

  (16) We **made** him [go to the store again]

- **VPto** (*To*-Marked Infinitive Verb Phrase)

  (17) What should she **do** [to test her hypothesis]?

- **VPed** (Participial Verb Phrase)

  (18) The twist it included in the storyline **had** me [tickled].

- **VPing** (Gerundive Verb Phrase)

  (19) Visitors don't **enjoy** [filling out HTML forms].

4.1.4 Clause Types

4.1.4.1 Finite Clause

- **Sfin** (Finite Clause (with or without *that*))

  (20) Last night I **learned** [that surrealism isn't just a Salvador Dali thing].

- **Sinterrog-** (*Wh*-Clause)

  (21) Could you **tell** me [how the Hawaiian Islands formed]?

- **Swhether** (*Whether/if*-Clause)

  (22) She told the BBC she did not **know** [whether the man who survived 7.2 parts per thousand had set a world record].

4.1.4.2 Nonfinite Clause

- **Sing** (Gerundive Clause)

  (23) My mom doesn't **like** [me being a vegetarian]!

- **Sto** (*To*-marked Clause)

  (24) I'd **like** [you to say hi to my sister].

- **Sforto** (*For-to*-marked Clause)

  (25) In 1937, a friend **arranged** [for Reagan to take a screen test]

- **Sbrst** (Bare Stem Clause)

  (26) Deborah **requests** [that she be allowed to live in a town nearby].

4.1.5 Subordinate Clause (Sub) with subordinating conjunction

(27) [Although Smithers credits Lightnin' Hopkins and Mississippi John Hurt as early influences], he **says** he's not really a bluesman in the classic sense.

(28) Alex **considers** Smithers to be one of his best friends [because Smithers is always looking out for him and making sure he's safe].

4.1.6 Adjective Phrase Types

- **Non-maximal Adjective (A)**

  This phrase type label is used for relational adjectives modifying noun targets(cf. section 3.4.2 on relational modification and compounds).

  (29) [economic] **policy**
  (30) [educational] **excellence**

- **Standard Adjective Phrase (AJP)**

  (31) Philip has [bright green] **eyes**.
  (32) The light **turned** [red].

4.1.7 Adverb Phrase (AVP)

(33) All items at [greatly] **reduced** prices!

(34) I've been doing that all night and, [quite frankly], my jaw **aches**.

4.1.8 Quantifier Phrases (QUANT)

(35) I have [two] **bottles** of correction fluid on the stand beside my favorite seat.

4.1.9 Quote (QUO)

(36) ["Could his performance tonight make or break the campaign?"] **exhaled** John Gibson of Fox News Channel.

(37) ["Hush, dear,"] Ruth **whispered**, ["I know, and I'll tell you some time, but I don't want her to know."]

4.2 Phrase Type Labels for Noun Phrases

4.2.1 Non-referential NPs

The distinction between **referential** and **non-referential** NPs warrants attention. Expletive *it* and *there* are the two kinds of non-referential NPs. As pointed out above, these constituents are given neither FE labels nor GF and PT tags, although we record their presence on the Other Layer. In such cases, FrameNet syntactic tagging cannot be directly mapped onto ordinary syntactic parses.

Some examples are given below.

(38) [It] is **clear** that we won't finish on time.

(39) [It] is **odd** that George is winning.

(40) In the same year [there] **arrived** from France the Rev. Louis Maigret.

(41) [There] **ensued** a great controversy, as some of the users didn't particularly want to be called losers to their faces every time they used the computer.

4.2.2 Possessive Noun Phrase (Poss)

Referential NPs are either possessive NPs (marked **Poss**) or standard (non-possessive) NPs (marked **NP**). Possessive NPs, which may either be possessive pronouns or noun phrases marked with *'s*, often express frame elements of predicating nouns. For example, in the Communication frame, possessive nouns express the SPEAKER role when they are the determiners of target nouns such as *claim*, *remark*, *reply*, etc.:

(42) I question [your] **claim** that the car was already damaged.

(43) [The President's] **remarks** surprised the reporters.

(44) [Leslie's] **reply** was well-timed.

The phrase type Poss is always paired with the grammatical function Gen(itive).

Note: The label Poss is not restricted to NPs denoting actual possessors. It is a morphosyntactic type rather than a semantic type. In this connection, note that *of*-PPs whether or not they denote Possessors never get the phrase type Poss. It should further be noted that given our lexicographic purposes, there is no reason to have a special category covering the complex NP type "a remark of the President's" or "a friend of mine". That is, we believe no frame-bearing word will specifically identify such phrases among their valence members.

4.2.3 Non-maximal Nominal (N)

In some situations it is necessary to tag nominal expressions which are not complete (i.e. maximal) noun phrases. For example, consider nominal modifiers of target nouns, as in examples (45) and (46) below, or the modified nouns in sentences showing target adjectives used attributively, as in the second pair of examples, (47) and (48).

(45) The judge dismissed the [forgery] **allegations**.

(46) [Cancer] **treatments** are advancing rapidly.

(47) **Allergic** [patients] benefit from this medicine.

(48) The senator gave a **polemical** [acceptance speech].

These non-maximal nominal expressions are given the phrase type **N** (for 'nominal').

In contrast, head nouns that are frame elements of post-nominal modifiers are not treated as non-maximal nominals. Instead they are treated as if the post-nominal modifier was actually predicated of the nouns in a copular clause. Thus, they are labeled full NPs with respect to Phrase Type, and as External arguments with respect to Grammatical Function.

(49) The problem seems to affect [people NP/External] **sensitive** to primulas.

4.2.4 Standard Noun Phrase (NP)

We treat as standard Noun Phrases all nominals that are not excluded as non-referential noun phrases—recall that these latter are not assigned frame elements, phrase types, and grammatical functions at all—or assigned the phrase type labels possessive (Poss) and non-maximal nominal (N). Standard Noun Phrases are marked with the phrase type **NP**.

(50) I **heard** [an interesting story] today.

(51) I **dropped** [the lid] on my foot.

Since we have a commitment to tagging full constituents rather than only their headwords with frame element labels (cf. section 3.2), modifiers and complements are included in the tagged noun phrases, as shown below. Notice in particular, as shown by examples (52) to (57), that we include both restrictive and non-restrictive relative clauses as well as appositives in the NPs we tag.

(52) [The cat in the corner] **likes** celery.

(53) [The cat running down the hall] **has** the catnip.

(54) **Stop** [that cat with orange stripes running down the hall].

(55) [The cat that's sitting on your lap] sure **sheds** a lot of hair!

(56) [My father, who worked as a clown all his life,] **refuses** to laugh at any of my jokes.

(57) [My uncle, Hollingworth Bowler III,] likes to **tell** stories from his sea-faring days.

Note further that standard NPs do not have to be headed by nouns. We treat free relative clauses (also called headless relative clauses) as NPs also. Likewise, we consider bare numerals as complete NPs.

(58) Harry will **eat** [what Sally is eating].

(59) I **want** [two].

4.3 Phrase Type Labels for Prepositional Phrases

PP is assigned to ordinary prepositional phrases with nominal objects and to particles, the latter under the assumption that particles can be regarded as prepositional phrases which lack objects.

(60) The passengers **looked** [at the monitors].

(61) Please **put** the vase [down].

PPing is assigned to prepositional phrases with gerundial objects rather than nominal ones.[2] Here are some examples:

(62) The fog **prevented** us [from seeing anything].

(63) They **found** Bill [reading the newspaper as if nothing had happened].

In addition, we assign the phrase type PP to the second piece of certain discontinuous degree phrases, as exemplified below.

(64) Billy Jeidels is so **wicked** [as to be beyond redemption].

(65) Stealing paper clips from work is less **immoral** [than taking them from a store PP].

These are closely related to the *than-* and *as*-phrases in the following sentences, which are also tagged as PP.

(66) Claire is as **wicked** [as John PP].

(67) Ada is less **moral** [than Dave PP].

[2]See section 4.4.2.5 on verb phrase types for a discussion of how we distinguish between verbal *-ing* forms tagged as VPing and nominal ones that are treated as NPs. This distinction applies to the case of prepositions governing *-ing* forms: some cases are treated as PPing and others as PP.

The PT-label **PPinterrog** is used for prepositions that govern finite or non-finite *wh*-clauses and verb phrases.

(68) That **depends** [on who did it PPinterrog]

(69) He **asked** [about how she was doing PPinterrog]

(70) Are you **confused** [about when to start meds PPinterrog]?

4.3.1 Particles

Particles like those in the following examples are treated as prepositions without objects and are assigned the Phrase Type **PP**.

(71) Did the police **escort** the streaker [out]?

(72) He **put** his hand [in] and felt for the ring.

Note that particles of this kind may occur before NPs and therefore give the appearance of being the heads of regular PPs with NP objects.

(73) Did the police **escort** [out] the streaker?

(74) He **put** [in] his hand and felt for the ring.

However, given that they are separable, as shown in the earlier examples, they cannot plausibly be treated as the heads of PPs in these contexts: *out the streaker* and *in his hand* are not constituents in (73) and (74), respectively. Therefore, they are assigned a separate label **PP**.

Whether or not a word W is to be treated as a particle can be established by this criterion: if verb V and particle W can be used as acceptable paraphrases of each other both in verb phrases of the form *V W NP* and of the form *V NP W*, then the word W is a particle rather than a preposition governing an NP.

While some particles, like *in* in (74) above, are equivalent in form to prepositions and may therefore misleadingly appear to head PPs in certain contexts, other particles do not resemble prepositions and are therefore less likely to be mis-analyzed that way:

(75) The referee **kicked** [away] the ball just as I was getting ready to take the free-kick.

(76) The librarian told me to **put** [back] the books.

There is no syntactic reason to distinguish the particles in (75) and (76) from the ones which resemble prepositions, and they are therefore given the same label (PP).

Finally, note that a verb-particle combination may be either a productively formed combination or a multiword expression (multilexeme lemmas). What is crucial in this connection is whether the verb could have the same meaning when the particle is either missing or when it is replaced by a different particle or preposition. *Take off*, for instance, is found in the Undressing frame; *take up* in the Activity_start frame; and *think up* in the Invention frame. In all these cases, the particle cannot be omitted with the frame-appropriate meaning intact. These combinations are thus to be treated as multiword expressions where no FE/GF/PT triple is assigned to the particle since it is part of the target and accordingly marked with the Target label on the Target-layer.

4.3.2 Prepositional Verbs

Some verb-preposition combinations are clearly conventional, as shown here.

(77) The passengers *looked at* the information monitors.

(78) Let me know if you *come across* that reference I asked you about.

We analyze the prepositions in such expressions as heading PPs.

(79) The passengers **looked** [at the information monitors].

(80) Let me know if you **come** [across that reference].

Though these verb-preposition combinations are units in the lexicon, we do not capture their unitary status in terms of *constituent structure*. That is, we do not analyze *look at* and *come across* as syntactic constituents.

In accordance with the Construction Grammar analysis of these expressions, their unitary status is captured in the valence representations of lexical entries. For example, there will be a lexical entry for *look at* which states that the verbal head *look* requires a PP headed by the preposition *at*.

4.3.3 Complex Prepositions

Some prepositions function as individual lexical units relative to a target verb but orthographically consist of more than one word (complex prepositions shown in italics).[3]

(81) **Put** the birthday cake *next to* the other desserts.

(82) He **mowed** the lawn *instead of* me.

Expressions of this kind are treated as single complex prepositions which head normal PPs. The PPs in the above sentences should be tagged in the following way:

(83) Your birthday cake was **put** [*next to* the other desserts].

(84) He **mowed** the lawn [*instead of* me].

4.3.4 Preposition Stranding

A preposition and its complement may be separated from each other, with the preposition appearing in a canonical post-verbal position and its complement noun phrase appearing pre-verbally in the clause, in a position that is not syntactically licensed by the verb.

(85) [John] we **laughed** [at].

(86) [Who] did you **refer** her [to]?

Since allowing for preposition stranding is not lexically relevant information, FN avoids annotating such sentences. If, however, sentences with preposition stranding have been annotated, then the two parts are assigned the same phrase type value, namely PP. (The two pieces also share the same grammatical function value, Dep.)

4.3.5 Prepositional Phrases with Relative Clauses

If the target word is inside the relative clause, we mark the prepositional phrase containing the relative pronoun or relativizer as a frame element as usual. The noun phrase antecedent to the relative pronoun or relativizer receives the identical FE/GF/PT triple. Ant and Rel labels are applied on the Other layer to the antecedent and the relative pronoun, respectively, as shown earlier.

(87) [the house Ant] [out of [which Rel]] I was **evicted**

(88) [the operator Ant] [to [whom Rel]] he had **spoken**

Notice that both in (87) and (88) the second bracketed constituent is treated as a PP.

If preposition stranding occurs within the relative clause, we proceed in the way described earlier in section 3.2.6. The stranded preposition carries a frame element label and shares its phrase type (PP) and grammatical function (Dep) both with the relative phrase (if there is one) and with the antecedent. The Antecedent and any relative word present carry the Ant and Rel labels on the Other layer. The stranded preposition, of course, carries neither of those labels, as is shown in Figure 4.2.

[3]Note, however, that when focusing on the prepositions as targets of analysis, we posit a preposition *next to*, that is a two-lexeme lemma, and a simple preposition *instead*. The different treatment is motivated by the fact that *instead* can be used without an overt *of*-complement (expressing the non-actual role filler) whereas *next* occurs in this meaning only together with *to*.

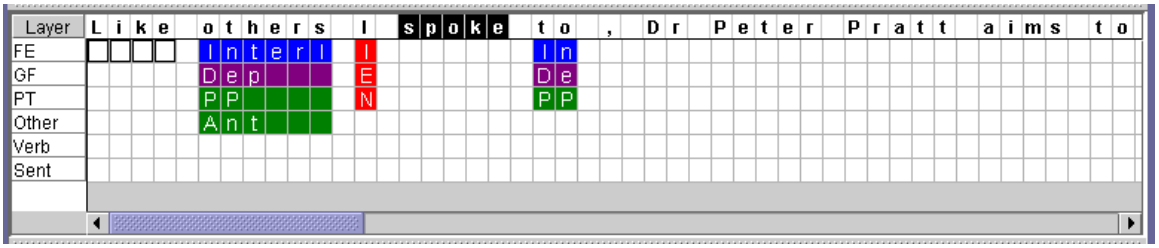

Figure 4.2: A relative clause with preposition stranding (repeated from above)

4.4 Phrase Type Labels for Verb Phrases

Every verb phrase has at least a head verb, which may be a main verb or an auxiliary. VPs headed by main verbs may also contain one or more auxiliaries. A verb phrase may also have a negative marker, an infinitive marker, a pre-verbal adverb phrase, one or more complements of the verb, and one or more post-verbal adjuncts. A VP may be headed by the main verb in a sentence or it may be embedded as a complement under another verb. The following examples show a variety of VPs, where the VPs are italicized.

(89) I *have*. (In response to "Have you taken out the trash?")

(90) This book *really stinks*.

(91) I didn't expect you *to eat your sandwich so quickly*.

4.4.1 Finite Verb Phrases (VPfin)

Any VP containing a verb (including auxiliaries) which (1) expresses information about tense and (2) is not in a separate embedded clause is tagged as a finite VP. Finite VPs are not generally subcategorized for, but it is nonetheless necessary to tag them in certain contexts, as illustrated here.

(92) Who do you **think** [ate the sandwich]?

(93) What did you **say** [fell on your hat]?

This pattern seems to be limited to a fairly small number of verbs of belief and assertion which subcategorize for clausal complements: *think*, *believe*, *say*, *claim*, *assert*, etc.

In contrast to the irregular verbs *eat* and *fall* in the above examples, note that finite past tense verb forms are frequently identical in form to past participial forms, e.g. *played-played*; *stated-stated*; etc. The participles are, however, non-finite and they are not covered by the tag VPfin (cf. section 4.4.2.4).

4.4.2 Non-finite Verb Phrases

Among non-finite VPs, it is necessary to recognize bare stem infinitives (**VPbrst**), *to*-marked infinitives (**VPto**), verb phrase relatives marked by *to* (**VPtorel**), past participial phrases (**VPed**), and gerunds (**VPing**).

4.4.2.1 Bare stem infinitives (VPbrst)

Bare stem infinitives are non-tensed verb phrases headed by verbs in the bare stem form without the infinitive marker *to*. Examples of bare stem infinitives (VPbrst) are given below.

(94) We **made** the children [take naps].

(95) Management **let** the employees [set their own hours].

Note that *the children take naps* and *the employees set their own hours* are not treated as so-called small clauses in the FrameNet project, though that is how they are sometimes analyzed.

4.4.2.2 *To*-marked Infinitives (VPto)

To-marked infinitives are VPs that begin with the infinitive marker *to*. Otherwise they are identical to bare-stem infinitives. Examples of *to*-marked infinitives appear below.

(96) The cat **wants** [to go outside].

(97) The mayors **persuaded** the President [to support the cities].

(98) It is **hard** for children [to tie their own shoes].

(99) I wish John Edward all the luck in the world and hope there is some element of truth in his **claim** [to be able to speak to the dead].

Note that *to*-marked infinitives that occur with noun targets are not always instances of VPto as in (99) above. Some of them serve as restrictive relative clauses and are then given the phrase type VPtorel, as discussed in section 4.4.2.3.

4.4.2.3 Verb Phrase Relatives(VPtorel)

Noun-modifying non-finite relative clauses headed by a *to*-marked VP are marked as VPtorel, regardless of whether they include a *wh*-word or not. (There is a separate phrase type label *Srel* for finite relative clauses.)

(100) **Towels** [to dry yourself with] can be found in the closet on the left.

(101) Nietzsche insists that there are no rules for human life, no absolute values, no **certainties** [on which to rely].

As a relative clause type, the VPtorel phrase type is, of course, restricted to dependents of noun targets.[4]

Notice that when verb phrase relatives (and also clausal relatives) are annotated relative to a head noun target, no marking of Ant and Rel on the Other layer takes place.

4.4.2.4 Participial Verb Phrases (VPed)

Participial Verb Phrases are VPs that begin with a past participle, typically ending in *-ed*. These phrases usually occur as post-nominal modifiers of target nouns, as in (102) and (103), but are also found as complements of certain control verb targets such as *have*, as in (104).

(102) The witness believes that the **man** [shown on the photograph] is the bank robber.

(103) By discussing the **events** [covered in the news], parents can help their children gain a better understanding of the world in which they live.

(104) The pastor interrupted the service in the middle of his sermon and **had** the man [forcibly removed].

4.4.2.5 Gerundive Verb Phrases (VPing)

Gerundive VPs are VPs headed by verbs in the *-ing* form. They often occur in syntactic contexts in which nominal expressions also occur. Examples of Gerundive VPs are provided here.

(105) My friend **likes** [running barefoot].

(106) [Inhaling pepper] **makes** most people sneeze.

(107) We **watched** the dogs [playing].

Gerunds present a challenge because they are sometimes verb-like and sometimes noun-like. FN annotators consider both syntactic and semantic criteria to determine if the automatic classification of gerundive verb phrases is correct. In particular, if the *-ing* form takes the same arguments as the related verb, e.g. if it takes an object or is modified by an adverb, we tag it as VPing rather than NP, as in (106) above or as in *He **kept** [singing the Albanian national anthem].* Also, if the context makes it clear that the *-ing* form refers specifically to an action, we use the VPing tag: *We were **thinking** [about dancing tonight], but I twisted my*

[4] These dependents should be extra-thematic frame elements such as Descriptor; there are no nouns that lexically require modification by a relative clause of any type. The phrase type VPto discussed above, by contrast, is typically assigned to core and peripheral frame elements and occurs with targets from the three major lexical classes (verbs, adjectives, nouns).

ankle. In this example one could not substitute the noun *dance* and keep the same meaning. By contrast, *Let's* ***discuss*** *[dancing]* is adequately re-phrased as *Let's* ***discuss*** *dance* and one can therefore tag *dancing* as an NP. In addition, any *-ing* form that is determined like a noun by, for instance, *the, this, that*, or *a* (*the killing*), or modified by a possessive (*my dancing*), an adjective (*quick thinking*), or by a following *of*-PP (*sounding of the alarm*) is treated as a noun.

4.5 Phrase Type Labels for Clauses

FrameNet syntax treats certain expression types as combinations of smaller constituents in contrast to some syntactic theories which treat them as clauses. For example, the sequence *Pat leave* in a sentence like *They* ***made*** *Pat leave* is sometimes analyzed as a 'small clause,' but in the FrameNet metalanguage it is treated simply as an NP followed by a bare stem infinitive VP.

This strategy has been adopted for two reasons. First, it simplifies the lexicographers' task of annotation, making it unnecessary to decide in certain cases which combinations of constituents should be treated as clausal and which should not. Second, it makes the lexical descriptions produced by FrameNet relatively theory-neutral. While the question of which verbal complements are clausal and which are not is answered differently in different syntactic theories, the analysis of clauses into their major constituents is uncontroversial in most cases.

As the reader will notice, there sometimes is no parallel between verb phrase types and clausal phrase types. This is true, for instance, for phrase types that figure in main clause and embedded questions. The label *Sinterrog* (cf. 4.5.3) covers both finite clauses and non-finite verb phrases because there are no predicates that specifically select either finite *wh*-clauses or non-finite verb phrases. Another case where there is a lack of parallelism involves 'small clauses'. Small clauses that are arguments of a target predicate are divided up into an NP and a secondary predicate, except for cases tagged as *Sing* (cf. 4.5.5). By contrast, small clauses that modify NPs or clauses are assigned to the single category *Sabs*. (These modifying small clauses are said to figure in absolutive constructions, hence the name *Sabs*, and they are typically tagged as the extra-thematic frame elements Depictive or Event_depictive.) Not all kinds of small clauses that can appear as arguments can also appear in absolutive constructions: compare **[Bill to arrive], John hid the money* and *I want [Bill to arrive]*. Thus, we lose some formal information by not recording the specific subtypes of absolutive clauses that occur in the data. However, the form of an absolutive construction is not lexically selected, in distinction to the kinds of 'small clause' that a predicate can take as an argument: *I saw him leave* v. **I saw him to leave*. From a lexicographic point of view, our treatment is therefore adequate.

4.5.1 Absolutive non-finite clauses

In particular for the annotation of Depictive FEs, we introduced a PT Sabs for small clauses modifying either a participant of the main clause as in (108)-(109) or the frame instance evoked by the main clause predicate (110).

(108) [His hands in his pockets Sabs], he **shuffled** back out of the room to wait until Unca had time to talk to him again.

(109) Purring loudly , Cas **padded** towards her, [tail erect Sabs] , [bright green eyes unblinking Sabs].

(110) [Both sections smashed to flinders Sabs], he **could** not put it back together before Mom got home.

Note that this label also applies to absolutive constructions that are introduced by *with*. In cases such as (111), *with* functions as a subordinator rather than as a preposition: notice that the predicate *out of the window* is required, which would be unusual if the phrase were a normal *with*-PP.

(111) [With both feet sticking out of the window Sabs], she **evinced** great surprise when the officer pulled along side.

4.5.2 Declarative Finite Complement clauses (Sfin)

Declarative finite complement clauses are full sentences that may begin with the complement marker *that*. In this PT, the entire clause, including the complement marker, is tagged.

(112) Pat **knew** [Kim would never agree].

(113) Pat **knew** [that Kim would never agree].

4.5.3 *Wh*-interrogative Clauses (Sinterrog)

Structurally, a *wh*-interrogative clause may be a sentence or a verb phrase. Although not full clauses, the interrogative verb phrases only occur in constructions which allow a full SINTERROG as well, and therefore a single PT is used for both. Note that we treat *how* as a *wh*-expression. *Wh*-expressions are included in the tag for the clause.

(114) I **heard** [what you said].

(115) I **forgot** [what to say].

(116) I **know** [how you feel].

(117) I don't **know** [how to react].

(118) I **asked** [who came].

(119) She **told** me [who to invite].

4.5.4 *Whether-if* Interrogative Clauses (Swhether)

Structurally, a *Whether-if* interrogative clause may be a sentence or, in the case of *whether*, a verb phrase. Although not full clauses, these phrases only occur in constructions which allow a full *Whether-if* clause and therefore a single PT is used for both.

(120) I **wonder** [whether the Indian restaurant delivers].

(121) He **wondered** [whether to turn back].

(122) Kim didn't **know** [if Pat liked the show].

4.5.5 Gerundive Clauses (Sing)

With certain predicates, sequences of object-form noun phrase and gerundive verb phrase are treated as single clauses by FrameNet. The reason for the analysis as a clause is that with the predicates in question the noun phrase cannot be separated from the gerundive verb phrase, for instance, in passivization. Compare *like*, which takes an Sing phrase, to *see*, which takes an NP and a VPing phrase.

(123) I don't **like** [him being here all the time].

(124) [*He] wasn't **liked** [being there all the time]

versus:

(125) You could **see** [a muscle] [jumping in Hubert Molland's cheek].

(126) [A muscle] could be **seen** [jumping in Hubert Molland's cheek].

Note: gerundive forms with a possessive subject (e.g. his) which may look similar to Sing clauses are treated as noun phrases:

(127) I don't **like** [his being here all the time].

4.5.6 Finite relative clauses (Srel)

The label *Srel* is used for all finite relative clauses regardless of being introduced by a *wh*-word, *that*, or zero, and whether the clause is interpreted as restrictive or non-restrictive.

(128) The **guy** [who I bumped into on the train Srel.Dep] was Herbert Kornfeldt.

(129) The **guy** [that I bumped into on the train Srel.Dep] was Herbert Kornfeldt.

(130) The **guy** [I bumped into on the train Srel.Dep] was Herbert Kornfeldt.

(131) My **neighbor**, [who I keep bumping into], gives me angry looks these days.

It should be noted that the label Srel does not apply to so-called free or headless relative clauses such as *He lives off* [*what he can sell at the farmer's market*]. Such clauses are tagged with the phrase type label NP since they are distributed like noun phrases.

4.5.7 *To*-marked clauses (Sto)

(132) I'd **like** [you to meet my mother].

(133) Certainly, but I should **hate** [you to forget that he has scored more runs in Test cricket than any other Englishman].

In sentences like the above examples, *you* cannot be the subject of a passive and therefore is treated as part of the non-finite clause.

(134) *[You] would be **liked** [to meet my mother]

4.5.8 *For-to*-marked clauses (Sforto)

Some nouns and verbs take clauses consisting of a *for*-marked subject and a *to*-marked infinitival verb phrase.

(135) I'd **like** [for you to meet my mother].

(136) I would **prefer** [for John to stay in the 250 class].

(137) The problem you've got is your **claim** [for him to contribute retroactively to that expense].

For-to-clauses need to be distinguished from just accidental juxtapositions of *for*-PPs and to-marked verb phrases, as in (138).

(138) We can throw a **party** [for him] [to show him that we are his friends].

4.5.9 Bare stem clauses (Sbrst)

Some verbs and nouns, typically with a semantics involving ordering or commanding, take a clausal complement with the verbal head in the bare stem form, that is, it is identical to the infinitive but lacks marking with *to*.

(139) The manager **demanded** [that employees be on time].

(140) The conversation resulted in my **insistence** [that I be shown at once the place where Jones and the others had met their fate].

4.5.10 Subordinate Clauses (Sub)

Certain clauses introduced by subordinators can instantiate frame elements and consequently need to be tagged. Such clauses receive the phrase type value Sub (Subordinate Clause) rather than Sfin (finite complement clause). Typically, frame elements that are assigned the phrase type *Sub* have peripheral or extra-thematic status (e.g., TIME, REASON). Note that we do not distinguish between the lexical categories of predicates heading the subordinate clauses. That is, examples (141) to (144) all count equally as instances of Sub even though the subordinate clause is headed by a finite verb in (141); by a VPing in (142); by an adjective in (143); and by a preposition in (144).

(141) I **admire** her [because she is an actress who can also sing Sub.Dep].

(142) [When considering such abominations Sub.Dep], we must be **concerned** for our precious bodily fluids.

(143) [When not quite sober Sub.Dep], Mila **likes** to sleep on the floor.

(144) [Though of noble lineage Sub.Dep], the Count **liked** to work as a video store clerk.

Certain adjective targets welcome discontinuous Degree phrases, as in the following example, where we assign the phrase type **Sub** to the *than*-phrase.

(145) That wine is more **delectable** [than I could imagine Sub.Dep].

4.6 Phrase Type Labels for Adjective Phrases

Adjective Phrases typically occur as pre-nominal modifiers, as non-Subject complements of *be* and a small number of other copular verbs (*seem*, *become*, etc.), and as predicate complements of verbs like *find*, *consider*, etc.:

(146) They were eating [very large] **sandwiches**.

(147) The **house** is [empty].

(148) You **seem** [sad] today.

(149) The company **considers** these documents [extremely valuable].

4.6.1 Standard Adjective Phrase (AJP)

An Adjective Phrase may consist of just a single adjective, an adjective with some modifying expression (such as an adverb or an intensifier), or a conjunction of adjective phrases:

(150) We **found** the play [dull].

(151) We **found** the play [extremely dull].

(152) We **found** the play [extremely dull and too long].

4.6.2 Non-maximal Adjectival (A)

Some tagged adjectival expressions are not treated as complete (i.e. maximal) adjective phrases. This is typically the case with relational modification, which is semantically very similar to noun-noun compounding (cf. 3.4.2) in not allowing for degree modification, as shown by (153) and (154), and in not allowing for a paraphrase in which the modifier is predicated of the head, as shown by the unacceptable contrasts in (155) and (156).

(153) [marital] **bliss** v. *[very marital] **bliss**

(154) [fruit] **juice** v. *[very/high fruit] **juice**

(155)
a. The Party of the Institutionalized Revolution (PRI) devised a new [economic] **policy**.
b. *The **policy** that the Party of the Institutionalized Revolution (PRI) has devised is [economic].

(156)
a. I got stuck in [rush hour] **traffic**.
b. *The **traffic** I got stuck in was [rush hour].

Relational adjectival modifiers are given the phrase type A, paralleling the treatment of nominal modifiers of nouns, which, as non-maximal nouns, are assigned the phrase type N.

4.6.3 Adjectives with Complements

Some adjectives take complements other than the nouns they modify, and these are included as part of the Adjective Phrase.

(157) When did you first **become** [interested in dinosaurs?]

(158) I **got** [scared of the typing sound].

In addition to adjectives like *interested, fond, afraid, scared* etc., the comparative forms of gradable adjectives account for a large number of complement-taking adjectives.

(159) Leslie **got** [taller than Kim AJP.Dep].

(160) Bart **turned** [angrier than he'd felt in a very long time AJP.Dep].

An adjective and its complement may form a discontinuous constituent when they modify a head noun. This is very common with morphologically comparative adjectives, as in (161), but also occurs with other adjectives that semantically denote comparison, as in (162).

(161) We need to find a [taller AJP.Dep] **player** [than Kim PP.Dep].
(162) There's also a white Reignac, which has a [similar AJP.Dep] **bottle** [to the one described by Douglass PP.Dep]

In such cases, both the adjective and its complement have the same frame element label applied to them, and they are assigned the same grammatical function label. However, the two pieces do receive separate phrase type values.

4.7 Phrase Type Labels for Adverb Phrases (AVP)

Adverbs, too, may express frame elements of a target verb, as illustrated in examples (163) to (165). They are assigned the phrase type AVP.

(163) The President **answered** the question [affirmatively].

(164) Bill **phrased** the rejection letter [carefully].

(165) Henry **left** the building [quickly] when he heard the police siren.

Note that some words that at first glance may not appear to be adverbs are assigned the phrase type AVP. For instance, *home* as used in *He went home already* is treated as an adverb.

4.8 Phrase Type Labels for Quantifiers (QUANT)

When annotating in respect to target nouns, the preceding number or quantifier is given the phrase type QUANT. For example,

(166) Bob poured [two QUANT] **cups** of coffee.

(167) Bob poured [thirty seven QUANT] **cups** of coffee at the brunch.

(168) Bob drank [a QUANT] **glass** of wine.

Note that we treat cardinal numbers and *a* (= 1) in the same way.

4.9 Phrase Type Labels for Quotes (QUO)

Some verbs of communication take quoted material as a complement that is assigned the phrase type QUO (and the grammatical function Dep). For example:

(169) ["Get out of here!"] she **cried**.

(170) ["But, I, er, uh..."] he **stammered**.

(171) He **thought** to himself, [let em come, there was no turning back now].

Quoted material can be of any syntactic form, or syntactically ill-formed, for that matter. Because the distribution or 'external syntax' of quoted material does not depend on its internal syntactic structure, we use a separate phrase type to tag it. Only direct quotes as in (169)-(170) and interior monologue as in (171) are given the phrase type QUO. Indirect quotes always take the form of some other kind of specific phrase type, as shown here.

(172) They **asked** us [what we were doing there].(*Wh*-clause)

(173) The President **said** [that he would support the inner city].(*That*-clause)

Quoted material is easy to identify because it almost always appears in quotation marks, which is included inside the brackets marking the Quote constituent.

Sometimes quoted material forms a discontinuous constituent:

(174) ["Cities,"] he **said**, ["are a very high priority."]

In such cases, both portions of the quote should be assigned the PT QUO.

Chapter 5

Assigning Grammatical Functions

In FrameNet, we annotate example sentences from the point of view of one particular target word in a given sentence. Each constituent tagged with a frame element in respect to a target word is assigned a phrase type as well as a grammatical function tag in respect to that target. In fact, **only** constituents tagged with frame elements are assigned grammatical functions. While target words are occasionally tagged with frame elements, they are never assigned a grammatical function.

The grammatical functions (GFs) that are assigned describe the ways in which the constituents satisfy abstract grammatical requirements of the target word. FrameNet grammatical function labels do not describe surface-syntactic positions of the constituents to which we assign them. For example, suppose the following sentence is selected to exemplify grammatical properties of the target word *treat*:

(1) Circumstances forced the doctor to **treat** her enemies.

Although the word *circumstances* is the subject of the sentence as a whole, this fact is not of lexicographic interest (given the target *treat* and is not marked in any way in the example sentence. Instead, the NP *the doctor* is tagged as the external argument (Ext) of *treat*, even though it is not the surface subject of the sentence, because it satisfies a valence requirement of the verb *treat* outside the phrase headed by *treat* (thus 'external'). That is, it fills a semantic role associated with *treat*, which would be realized in a simple declarative main clause by the subject of the clause.

(2) The doctor also **treated** her enemies.

The set of grammatical functions that is commonly employed can be described in the following way. Constituents that occupy core syntactic slots fulfill the functions of Subject and Object. All other constituents accompanying a syntactic head are considered dependents given that their presence in a construction centered on the head is licensed by the head. Among dependents, one can further distinguish between obligatory complements and optional modifiers. Additionally, for nominal heads, a third category of dependents needs to be recognized, determiners.

As pointed out in section 3.2, FrameNet records a coreness status for each frame element and therefore has no need to encode the obligatoriness or optionality of a constituent with distinct grammatical function labels. Consequently, FrameNet does not have a pair of grammatical functions such as Modifier and Complement.[1] Obligatory and optional dependents that do not serve the core syntactic functions of External argument (subject) or Object both receive the grammatical function label Dependent.

It is worth reiterating that Modifier is used here in a purely syntactic sense to refer to optional dependents. Thus, all adjectives that pre-modify a target head noun and all adverbs that modify a target adjective or verb are assigned the grammatical function Dependent, even though, in a semantic sense, the term Modifier refers to the typical adjectival semantic function of supplying additional features of referents and the common adverbial semantic function of elaborating the scene evoked by a verbal or adjectival predicate. The semantic function of modification logically involves predication, which happens in the pragmatic background, rather than being pragmatically asserted. Accordingly, qualitative or intersective adjectives such as *red* in *red car* much more saliently take the head noun as a role filler rather than the other way around. As noted above, there are relational adjectives such as *economic* and *rural* that behave differently: they are basically role

[1] Before the marking of frame elements for coreness status was introduced, there was a Modifier-Complement distinction. It was subsequently collapsed and now only Dependent is in use.

fillers rather than frame evokers. From a semantic point of view, the grammatical function Dependent fits these relational adjectives more appropriately than it fits the intersective adjectives.

Although we generally do not distinguish syntactic Modifiers and Complements, there are two exceptions. First, FN has a separate grammatical function Appositive that is used with noun targets rather than using the general Dep label. Second, in the case of adjective and adverb targets we apply the label Head to the core semantic argument over which they predicate (***red*** *[house HEAD]*), thereby following the syntax of the overall noun phrase or clause rather than the semantics of the adjective or adverb, which would favor the use of the label Dependent (****red*** *[house DEP]*).

Note also that although we generally do not annotate Determiners of noun targets, we use the grammatical functions Gen(itive) for possessive determiners of nouns and Quant for quantificational pre-determiners of nouns when these elements fill frame element roles.

The combinations of grammatical function labels that occur with particular lexical items provide both a way of encoding the syntactic constructions a lexical item occurs in, and a way of retrieving certain subcases of grammatical functions that are not directly encoded by individual GF labels.

For example, the verb *like* can occur in sentences with extraposed objects:

(3) I **like** it [that you speak French].

FrameNet annotations of example sentences demonstrate that the verb *like* may take a null object *it* followed by a complement clause as one of its valence options. It is possible to retrieve examples of extraposed objects for *like* and other predicates from the database by searching for combinations of null object and complement clause.

5.1 List of Grammatical Functions

What follows is a list of the grammatical functions used in FrameNet. It is followed by sections providing detailed criteria for the assignment of each GF. As with phrase types, GFs are assigned automatically during the classifying process, but may require manual correction.

5.1.1 Grammatical Functions assigned by Target Verbs

External Argument (Ext)

Object (Obj)

Dependent (Dep)

5.1.2 Grammatical Functions assigned by Target Adjectives

External Argument (Ext)

Head noun modified by attributive adjective (Head)

Dependent (Dep)

5.1.3 Grammatical Functions assigned by Target Prepositions

External Argument (Ext)

Dependent (Dep)

5.1.4 Grammatical Functions assigned by Target Nouns

External Argument (Ext)

Genitive determiner (Gen)

Dependent (Dep)

Appositive (Appos)

5.2 Assigning GFs for Verbs

5.2.1 External Argument (Ext)

External covers situations in which phrases outside of the maximal phrase headed by the target word are linked functionally to the target word. This includes anything that satisfies an FE requirement of a target word in any of the following syntactic contexts, where the external argument appears in square brackets:

- as a subject of a finite target verb, as in (4)

  (4) [The physician] **performed** the surgery;

- any constituent which controls the subject of a target verb is assigned the GF External. This constituent might be a subject, as in (5), an object, as in (6), or a prepositional object of a governing verb, as in (7). (Governing verbs are rendered in typewriter font.)

  (5) [The doctor] `tried` to **cure** me.
  (6) They `persuaded` [the doctor] to **treat** me.
  (7) They `gestured` [to us] to **leave**.

- as a dependent of a governing noun, as in (8) to (11)

  (8) [The physician's] `decision` to **perform** the surgery....
  (9) Both of us have the utmost respect for [your] `decision` to **do** what you're doing.
  (10) We are glad for the [American] `decision` to **provide** relief.
  (11) Today's `decision` [by the Court] to **approve** our request for $1.1 billion in DIP financing is another positive step in our plan to emerge from Chapter 11.

Notice that the phrase *the physician's* is in the Genitive GF relation to the noun *decision*, but bears the Ext relation to the verb *perform*: like frame element labels, grammatical function labels are relative to particular target words.

By contrast, note that in example (7) the whole PP *to us* rather than just the NP *us* is tagged as the External Argument. Although we are only interested in the valence properties of *leave* and not in the valence properties of *gesture*, we have chosen to annotate the full dependent constituent of the controller which contains the frame element of the target word. The reason is that in cases like (9) it is not possible to annotate only an NP that can figure as a locally instantiated subject of the target.

5.2.2 Object (Obj)

Any normal object, any *wh*-extracted object, or any post-target-verb NP which controls the subject of a complement of the target verb is assigned the grammatical function Obj, as shown here.

(12) Voters **approved** [the stadium measure].
(13) [What] did you **cook** for dinner?
(14) They **expect** [us] to finish soon.
(15) They **made** [us] eat our vegetables.

In some syntactic theories, the NP *us* in examples (14) and (15) would be treated as the subjects of small clause complements, and not as Objects of the target words. In the context of FrameNet annotation, the grammatical function assigned depends on which predicate is chosen as the target of analysis. Relative to the finite control and raising verbs in (14) and (15), the post-verbal NPs are Objects, but relative to the target words *finish* and *eat* the two NPs in (14) and (15) are analyzed as having the grammatical function Ext.

The GF Object is also assigned to any subject of a *tough*-predicate which satisfies the Object role of a verb or preposition in the complement of the *tough*-predicate.

(16) [Artichokes] are hard to **eat**.

Generally, we do not assign the grammatical function Object to dependents of target verbs which cannot occur as the subjects of passive clauses. Typical cases of such post-verbal NPs are measure phrases as in (17) and (18).

(17) The boat **weighs** [12 tons].

(18) That book **cost** [$20].

There are several other kinds of contexts in which noun phrases that are not marked by a preposition can appear after a verb. Verbs like *ride, fly, take* which can be used to denote the use of a means of public transportation (as opposed to verbs like *drive* which denote the act of operating a vehicle) can appear with a bare NP denoting the means of transportation. Motion-related verbs such as *bike* in (21) can appear followed by bare NPs denoting an AREA that was significantly covered by the SELF_MOVER/THEME. We do not treat the noun phrases involved as Objects, but just as simple Dependents.

(19) My husband **flew** [United] last night and was served warm nuts as a snack.

(20) ??*United is **flown** by more than 5 million people a year.

(21) Sue and I **biked** [Europe] for our 1998 fund raiser.

(22) ??*Europe was **biked** (by us) for the 1998 fund raiser.

5.2.3 Dependent (Dep)

Dependent is the general grammatical function assigned to Adverbs, PPs, VPs, Clauses (and a small number of NPs) which occur after their governing verbs, adjectives or nouns in normal declarative sentences. This grammatical function includes both what are usually referred to as *arguments* and what are usually referred to as *adjuncts*. In the FrameNet annotation scheme, this distinction is captured via the coreness status assigned to a frame element and is not replicated in a grammatical function distinction.

5.2.3.1 PP Dependents

The GF Dependent is assigned to any particle or any PP, optional or obligatory, which expresses a semantic role belonging to the frame associated with the target word. Thus, setting adjuncts of PLACE (23) or TIME (24), purpose clauses (25), or other such expressions which can occur with very large classes of predicators still carry the same grammatical function Dependent as do core frame elements (26)-(28).

(23) I **ordered** a "steak" [in a St. Petersburg restaurant].
(24) [In 1790], New York City was **considered** the largest city in the United States with a population of 33000.
(25) Bill **sold** the house [in order to finance a concert].
(26) **Give** the gun [to the officer].
(27) Pat **spoke** [to me].
(28) Kim **phrased** the letter [with great care].

Some NPs are marked as Dependents rather than as Objects. These NPs are not passivizable, and they often express Place, Time, and other meanings normally associated with adjuncts and PP complements (e.g. Measurement). Such NPs are often Dependents in the same contexts in which comparable PPs might be used instead.

(29) I **run** [ten miles] every day.
(30) **Come** [this way]!
(31) I **expect** your papers [the moment you walk into class].

In keeping with conventions of Construction Grammar, the second object of ditransitives is treated as a Dependent. FrameNet does not have a category for second object and does not assign the grammatical function Object a second time in ditransitive sentences.

(32) They **gave** the children [candy].
(33) The children were **given** [candy].

A note on particles: Even if a preposition seems to 'go with' a verb to form a phrasal verb, if it can plausibly be considered the head of a post-verbal PP, it is analyzed that way, and the PP is assigned the GF Dep. Separable verb particles, like *up* in *pick up the package* (and *pick the package up*) are normally part of lexical units consisting of more than one lemma. As part of the target, they are not annotated with GF labels. The NPs that the particle verbs occur with (e.g. *garbage*) are treated as GF Obj. For more information on particles, see section 4.3.

In passive sentences, these NPs are treated as GF Ext.

(34) Pat **picked up** [the package Obj].
(35) Pat **picked** [the package Obj] **up**.
(36) [The package Ext] was **picked up** by Pat.

Locative adverbs such as *here* and *there*, whether used to express core or non-core frame elements, are normally assigned the GF Dep.

(37) Chuck **went** [there Dep].
(38) I feel like almost nobody **wears** hats [here Dep].

In some rare cases, however, *here* and *there*, like certain PPs, appear as subjects of predications in which some characterization of the location itself is given rather than it just serving as the setting for an event. Under such circumstances, they are assigned the GF label Ext.

(39) Understandably [here] is not the right **place** to report more in details.

(40) No wait, [over there] is even **better**.

(41) [Under the bed] is the best **place** to hide.

5.2.3.2 Verbal and Clausal Complements

Unless it serves as external argument, as in (42), any verbal or sentential complement, regardless of whether or not it is passivizable is assigned the GF Dependent, as shown here.

(42) [That it was once much greater] is **attested** by the many traces of former lazy-beds to be seen on the grazing land.

(43) They **want** [to stay home].

(44) They **expect** us [to stay home].

(45) I **believe** [that you are the winner].

(46) They **think** [you are the president].

(47) They **had** Peter [removed from the premises].

(48) I **wonder** [who will finish first].

5.3 Assigning GFs for Adjectives

5.3.1 External Arguments of adjectives

When an adjective appears in a clausal predication, one of its arguments is expressed as the subject of a copular verb (indicated with underlining) and of the clause as a whole, as in the following examples.

(49) [The chair] is **red**.
(50) [My sister] seems more **interesting** than yours.

This constituent is assigned the GF External Argument (Ext). Also, the noun in object-control constructions with adjectives is assigned the GF External Argument, illustrated below.

(51) We consider [Pat] very **intelligent**.

Note that copular verbs are not assigned GFs.

5.3.2 Modified head nouns with pre-nominal adjectives

In pre-nominal uses of adjectives, the modified head nouns are assigned the GF Head (Head).

(52) the **small** [children Head]

Only some types of sentences in which a target adjective appears in pre-nominal position are actually annotated and included in the database of corpus examples. These are the sentences in which the adjective has a *qualitative* as opposed to a *relational* use.

In a qualitative use, the modified noun expresses an element of the frame associated with the adjective, and this is the same frame element which is typically expressed by the subject of the verb BE or other clausal predication (in which the adjective occurs in predicate position), as shown below.

(53) The children are **small**.

Note that in these cases, annotators tag the verb with the label Copula on the (POS-specific) Adjective Layer.

In a relational use of an adjective, it is much more difficult to identify a specific frame element which is expressed by the modified noun. In fact, the semantic relation between the adjective and noun may be more strongly determined by the modified noun than by the adjective. In any event, the relation between modifier and head is much less tightly constrained than in cases of qualitative modification, resembling the relation between nouns in a noun-noun compound (cf. section 3.4.2). Very often relational uses of adjectives do not have corresponding predicative uses, as is apparent from the unacceptable example (55).

(54) Pat had an **immune** response to the virus.
(55) *Pat's response to the virus was **immune**.

We do not annotate relational uses of target adjectives because doing so is unlikely to reveal much that is interesting about the frames associated with the adjectives. Instead, they are captured as fillers of roles in frames evoked by the nouns they modify.

Many cases of relational modification (like many noun compounds) are highly conventionalized. Such cases are treated as idioms when they are relevant to a particular domain, and are identified with respect to their head nouns rather than their adjectives. For example, *immune response* is treated in the health care domain as a lexical unit headed by *response*, rather than as a particular use of *immune*.

5.3.3 Post-nominal adjectives

Arguments of post-nominal uses are treated as if the adjective were used with a copula. In particular, modified head nouns are assigned the grammatical function External (Ext) and the phrase type Noun Phrase (NP), rather than the grammatical function Head and the phrase type Noun (N). Thus, in the example (56), *people* has the grammatical function Ext and the phrase type NP. Note that in (57) the determiner *a* is included in the frame element in line with the analysis of the frame element as an external argument of *afraid*.

(56) The problem seems to affect [people] **sensitive** to primulas.

(57) She says she won't take orders from [a guy] **afraid** to come out of his cage.

5.3.4 Adverbial Modifiers of Adjectives

Adverbial modifiers of adjectives are assigned the GF Dep, regardless of whether the adjective modifies a noun or is used predicatively.

(58) The [very Dep] **beautiful** woman left the office.
(59) That was [so Dep] **silly** of you.
(60) [How Dep] **offensive** that movie was!

5.3.5 PP complements of Adjectives

Some adjectives require frame elements that are expressed by PPs, to which we assign the GF Dep.

(61) Smithers is not **aware** [of his bad behavior Dep].
(62) Lee is **certain** [of his innocence Dep].
(63) Bilger is **interested** [in the footwear of the Yeti Dep].

5.3.6 N complements of Adjectives

Some complement-taking adjectives that themselves pre-modify a head noun can be preceded by a second argument of theirs in a compound-like structure, as in (64).

(64) That day Meg and I were both suffering from [stupidity] **related** injuries.

Such examples look superficially like post-modifying uses of the target adjective but they are clearly different: the argument preceding the target has to be a bare nominal and cannot itself be determined. Compare (64) to (57) above in this regard.

5.4 Assigning GFs for Adverbs

Adverbial targets occur with only two different grammatical functions. The event or relation that they modify is tagged as Head and any element that modifies the target adverb receives the grammatical function Dep(endent).

(65) [Open the door Head] **carefully** so as not to bump it on the puppy's head.

(66) [That's Head] [quite Dep] **honestly** [a mistake on our part Head].

(67) **Quietly**, [Smith agreed to a new contract Head].

(68) We want [you to speak Head] **slowly**.

5.5 Assigning GFs for Prepositions

For most frames prepositions are not among the target words. However, they are targets in frames that cover the vocabulary of space, time, and motion. In these frames, we assign GFs for the relevant constituents.

5.5.1 Dependents of prepositions

Any constituent which occurs immediately after a preposition and which expresses a core element of the frame associated with the preposition is tagged as a Dependent (**Dep**). Typically this constituent is an NP, as in (69), but it can also be a gerund or a clause, as in (70) and (71), respectively. The different dependents of *before* in the examples below have different phrase types, as appropriate.

(69) We had a glass of wine **before** [the meal].
(70) Wash your hands **before** [returning to work].
(71) He left **before** [I had a chance to say anything].

The FrameNet analysis of post-preposition complements is different from other analyses in which at least NPs such as *the meal* in (69) would be called Objects of the preposition. FrameNet reserves the GF label Object for verbal complements.

The grammatical function DEPENDENT is also assigned to nominal or adverbial expressions that pre-modify certain types of prepositions, as shown in (72) and (73).

(72) The Control Room is [straight] **down** the corridor on your left.

(73) There is a parking sign about [10 meters] **past** the restaurant.

5.5.2 External Arguments of prepositions

A constituent which expresses an element of the frame associated with a preposition, but which is outside the PP, is tagged as an External Argument (**Ext**). The most easily identified prepositional External Arguments are those which occur with so-called reduced relative clauses (i.e. post-nominal modifiers) and with copular predications, as in (74) to (77).

(74) [the day] **before** yesterday
(75) [the trail] **to** our campsite
(76) [the ball] **under** the table
(77) [The ball] is **under** the table.

5.6 Assigning GFs for Nouns

5.6.1 Dependents of Nouns (Dep)

The GF Dep is assigned to any post-nominal complement of a target noun, examples of which are shown here.

(78) the **fact** [that cats have fur]
(79) a **letter** [to the President]
(80) a **story** [about a young entrepreneur]
(81) your **attempt** [to scare me]
(82) our **arrival** [at the station]

The GF DEP is also assigned to pre-nominal phrases of any type (noun, adjective, gerund, or participle) that fill frame element roles. Note that we make no distinction between cases such as (83) and (84), where the frame relation between the target noun and the pre-nominal phrase depends on the target noun, and others, such as (85) and (86), where it is evoked by the non-head modifier.

(83) [allergy] **treatment**
(84) [economic] **policy**
(85) [sleeping] **cat**
(86) [broken] **lamp**

Sometimes dependents of nouns are realized as the predicates of copular sentences, as in examples (87)-(89).

(87) The **fact** is [that cats have fur].
(88) The **letter** was [to the President].
(89) The **attempt** was [to scare you].

The copular verb that follows the target and introduces the 'predicate complement' in these expressions is tagged on the Noun layer with the Copula label.

5.6.2 External arguments of Nouns (Ext)

There are a few cases in which frame-evoking nouns have an External argument of their own. One is with support verbs, as in (90) and (91).

(90) [He EXT] made a **statement** to the press.
(91) [I EXT] took a **bath** for the first time in months.

Another is when the frame-bearing noun is governed by a control noun:

(92) [My EXT] attempt at an **agreement** with Pat failed.

5.6.3 Genitive determiner of noun (Gen)

The GF Gen is assigned to any possessive NP functioning as determiner of a target noun, as indicated here.

(93) [your GEN] **book**
(94) [your work's GEN] **influence** on the field

Note that the term Possessive (Poss) refers to the phrase type of Genitive NPs.

5.6.4 Quantificational determiners (Quant)

The GF Quant is assigned to a pre-nominal determiner of a target noun, including the indefinite determiner when it functions as a number, illustrated below.

(95) [two] **cups** of coffee
(96) [a] **glass** of wine

5.6.5 Appositives (Appos)

The GF Appositive is assigned to post-target appositional Ns or NPs.

(97) Libel **lawyer** [Jonathan Crystal APPOSITIVE] represented the plaintiff.
(98) Actor Robert Downey Jr. will walk down the aisle next year with **girlfriend** [Susan Levin APPOSITIVE].

Chapter 6

Semantic Relations and Types

The frames that we create, and thus the Frame Elements and Lexical Units associated with them, are intended to be situated in semantic space by means of frame-to-frame relations and semantic types. The relations we use include **Inheritance**, **Subframe**, **Causative_of**, **Inchoative_of**, and **Using**. There are several benefits to extensive use of these relations:

- Improving the comprehensibility of frames: The intended meaning of a more complex frame can be clarified by relating it to an existing, easily-understood frame.
- Robustness: In some cases, it would clearly be possible for other researchers (or ourselves) to have made different frame divisions than the ones we have made. Having relations to semantically similar frames allows frames (and thus their Lexical Units) to be associated despite being separated.

A number of other benefits are outlined in 6.3.

We have spent considerable time recently improving the semantic relations encoded in our data, and also making the relations more accessible via the FrameGrapher (see `http://framenet.icsi.berkeley.edu/\~{}FrameGrapher/`). We have also marked a large number of Frame Elements with semantic types indicating appropriate fillers (6.2.1). There remain a dwindling number of completely disconnected frames, all of which are semantically quite distinct from other frames that we already describe.

In addition, we have added two new types of relations, Precedes (replacing the meta-relations for Subframes) and Perspective_on, which replaces a subset of the Using relations with a more specific and informative relation. These should make our data more useful for all forms of computational processing. All of the frame relations referred to in this chapter *other than these new ones* can be visualized with the FrameGrapher on the main FrameNet website (`http://\~{}framenet.icsi.berkeley.edu`).

In the sections that follow, general descriptions of each of the Frame-to-frame relations (6.1) and semantic types (6.2) are given, and then a more formal description of the relations and the ways that they may be used for reasoning is discussed (6.3). Note that the following does not discuss the FE-to-FE relations, which have a closer relationship with valence and annotation than with semantically-defined relations. For these relations, see sections 3.2.2.1-3.2.2.3.

6.1 Frame-to-frame Relations

Each frame relation in the FrameNet data is a directed (asymmetric) relation between two frames, where one frame (the less dependent, or more abstract) can be called the Super_frame and another (the more dependent, or less abstract) can be called the Sub_frame. We give a more specific, informative name to the Super and Sub_frames for each relation, as shown in Table 6.1 below.

In general, each frame has one relation to some other frame, but there are occasional exceptions, as seen in the relations between Assistance and Intentionally_act, shown in Figure 6.1 below. Assistance inherits from Intentionally_act, with the HELPER bound to the AGENT of Intentionally_act, but Assistance also uses Intentionally_act since there is a second intentional action presupposed, namely that of the CO_AGENT, which is thus also bound to the AGENT role of Intentionally_act in a separate relation.

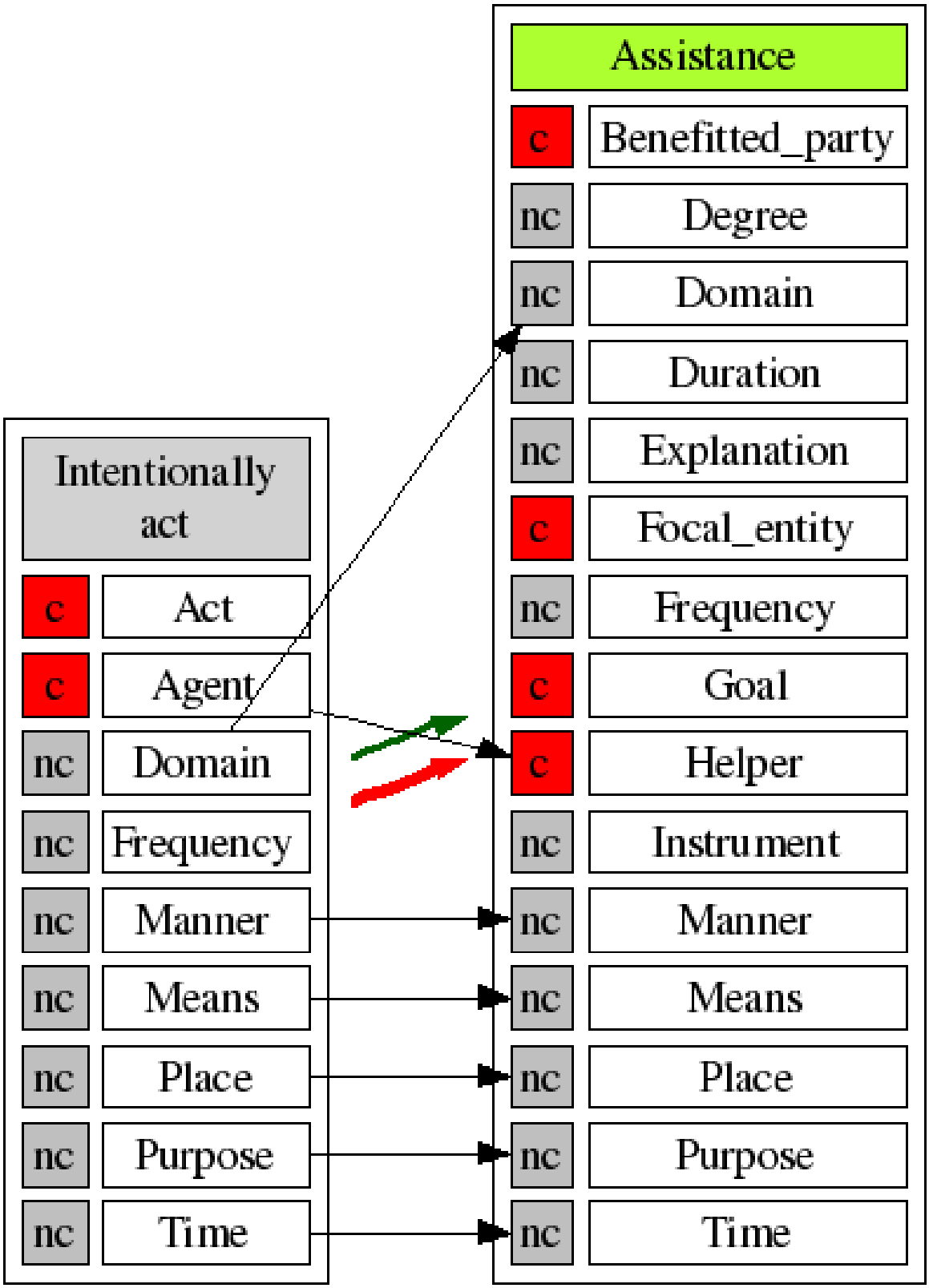

Figure 6.1: Assistance, with two relations to Intentional_act

Relation	Sub	Super
Inheritance	Child	Parent
Perspective_on	Perspectivized	Neutral
Subframe	Component	Complex
Precedes	Later	Earlier
Inchoative_of	Inchoative	State
Causative_of	Causative	Inchoative/State
Using	Child	Parent
See_also [1]	Referring Entry	Main Entry

Table 6.1: Types of Frame-frame Relations

6.1.1 Inheritance

This is the strongest relation between frames, corresponding to is-a in many ontologies.[2] With this relation, anything which is strictly true about the semantics of the Parent must correspond to an equally or more specific fact about the Child. This includes Frame Element membership of the frames (except for Extrathematic FEs), most Semantic Types, frame relations to other frames, relationships among the Frame Elements, and Semantic Types on the Frame Elements.

Properties of the Parent which are not strictly semantic in nature, such as not being evoked by lexical units (i.e. the Semantic Type Non-lexical_frame), being evoked by a particular set of Lexical Units, or having a See_also relation to another frame, are not inherited.[3]

Also, when there is a Core-set or an Excludes relation among Frame Elements of the Parent (see sections 3.2.2.1 and 3.2.2.3), these constitute disjunctive properties of the Parent. The Child frame may legitimately inherit only a subset of these disjunctive Frame Elements.

6.1.2 Perspective_on

This relation (new in Release 1.3) is a refinement of the more general Using relation (see 6.1.6). Perspective_on constrains related frames considerably more. The use of this relation indicates the presence of at least two different points-of-view that can be taken on the Neutral frame. For example, the Measure_scenario, in which an ENTITY's VALUE for some ATTRIBUTE is described, can be viewed either from the point-of-view of exact measurement (e.g. "Joey weighed 7 **pounds**.") or as a relative measure (e.g. "Joey was **heavy**."). The FEs in the two cases are quite different, so the words should not be included in the same frame (see 2), but they do make reference to the same scene. The Perspective_on relation allows us to refer directly to the scene (encoded by the Neutral Measure_scenario frame here) and connect the two. As in this case, the Neutral frame is normally Non-lexical and Non-perspectivalized. (See 6.2.2.)

A single Neutral frame generally has at least two Perspectivalized frames, but in some cases, words of the Neutral frame are consistent with multiple different points-of-view while the Perspectivalized frame is consistent with only one.[4] Whenever there is a state of affairs that is describable by a frame in a Perspective_on relation, all the other frames that are connected to it by the Perspective_on relation can also be used to describe the state of affairs.[5]

[2]In some few cases (enumerated in the release notes for data release 1.3), we have used the Inheritance relation in an apparently incorrect sense, where a Child frame has two types of realization, only one of which can straightforwardly be described as inheriting from the Parent. For example, the Manipulation frame has two possible points-of-view: one which focuses on the AGENT ("She **grabbed** the cigar.") and the other on the BODYPART_OF_AGENT ("Her hand **grabbed** the cigar."). Of these, only the AGENT-focused realization properly inherits from Intentionally_affect, as the BODYPART_OF_AGENT-focused realization has no AGENT, which is an obligatory FE of the Intentionally_affect frame. All the remaining errors in Inheritance are of the same kind: where the Child frame has an alternation of FEs related by some metonymy and one of these related FEs is not present in the Parent. Although not indicated in Release 1.3, in future releases these metonymies will be explicitly marked.

[3]These descriptions of a frame actually represent meta-information rather than true properties of the frame.

[4]Many of our frames, including many which are not explicitly involved in the Perspective_on relation, have more than one point of view inherently. In particular, any frames which have exclusion sets among their FEs have a separate point-of-view associated with each different choice of excluded FEs.

[5]Note that the fact that all of the frames in the set can describe a situation does not always mean that they are inter-substitutable.

Ex: These skateboards are easy to *sell*. ≠ ...easy to *buy*.

A word like *easy* is specifically sensitive to the point-of-view of its complement, picking out the actor in that scene as the EXPERIENCER. In a scene like the one involving *buy* and *sell*, there are multiple individuals that can be construed as the actor, and *buy* and *sell* lexically (and framally) profile these different potential actors. It is still the case that to whatever degree a buying event occurs, a selling event does also.

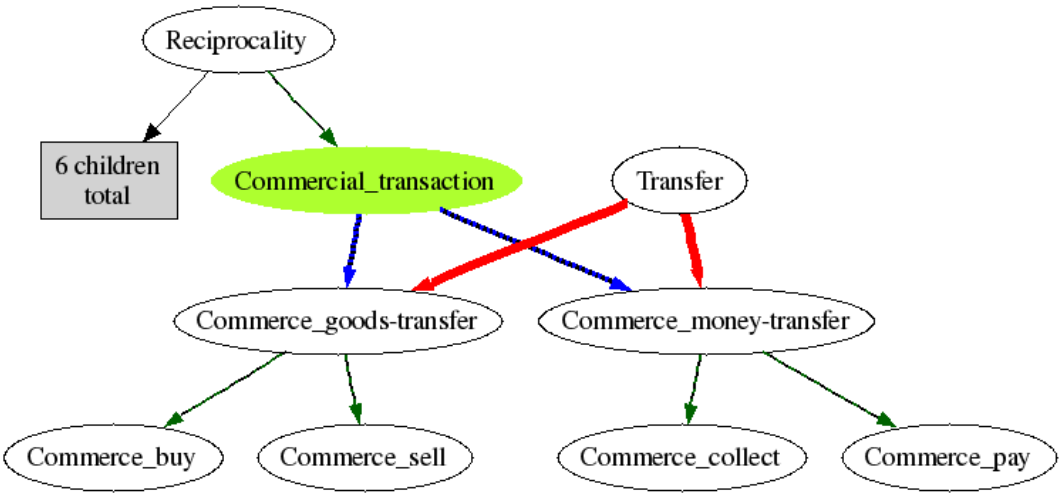

Figure 6.2: Commercial_Transaction

For example, the Commercial_transaction frame, diagrammed in Figure 6.2, specifies a complex schema involving an exchange of multiple Themes (the MONEY and GOODS) between the BUYER and SELLER, including also two subframes: Commerce_goods-transfer and Commerce_money-transfer. The Buying frame has a Perspective_on relationship with the Commerce_goods-transfer subframe in which the MONEY, GOODS, BUYER, and SELLER are identified.

Some other relations of this type include:

- Get_a_job and Hiring are perspectives of Begin_employment
- Rope_manipulation and Knot_creation are perspectives on Knot_creation_scenario

Further note that it is quite common for a frame to inherit from one frame and be a perspective on another. An example of this is again provided by the Commerce_buy frame, which inherits the Getting frame (not shown in the diagram 6.2) but is a perspective on the Commerce_goods_transfer frame. An act of buying is a sub-type of getting, which justifies the inheritance relation to Getting. Buying is an event that only occurs when the situation (the Commerce_goods_transfer) is also describable as selling. For that reason, it (and the Commerce_sell) is connected to the Commerce_goods_transfer frame via a Perspective_on relation.

6.1.3 SubFrames

Some frames are complex in that they refer to sequences of states and transitions, each of which can itself be separately described as a frame. The separate frames (called subframes) are related to the complex frames via the SubFrame relation. In such cases, frame elements of the complex frame may be identified (mapped) to the frame elements of the subparts, although not all frame elements of one need have any relation to the other. (In this respect, it contrasts with inheritance; see below.) Also, the ordering and other temporal relationships of the subframes can be specified using binary precedence relations. To illustrate, consider the complex Criminal_process frame, which is defined as below and whose frame relations are shown in Figure 6.3.

> A Suspect is arrested by an AUTHORITY on certain CHARGES, then is arraigned as a DEFENDANT. If at any time the DEFENDANT pleads guilty, then the DEFENDANT is sentenced, otherwise the DEFENDANT first goes to trial. If the FINDING after the trial is guilty, then the DEFENDANT is sentenced. In the end, the DEFENDANT is either released or is given a SENTENCE by a JUDGE at the sentencing.

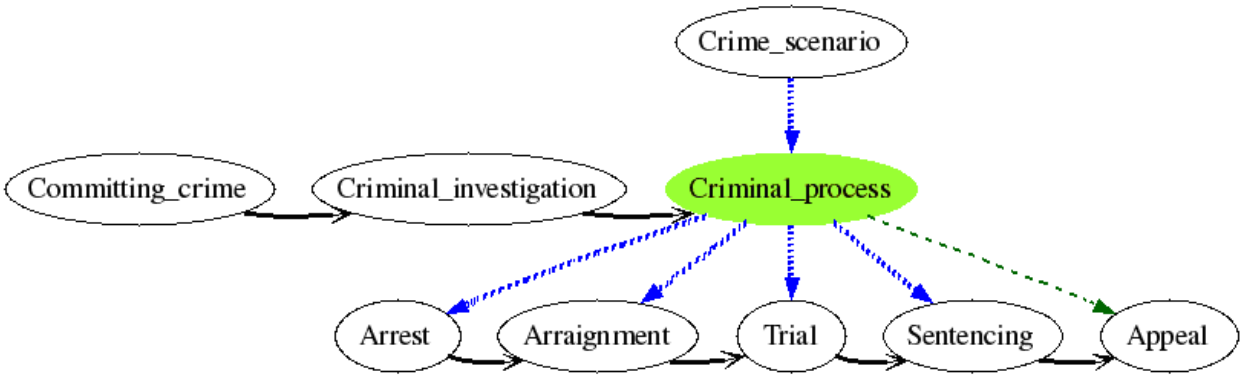

Figure 6.3: Subframes of the Criminal_process frame

For each step in the process, there is a separate frame in the database, including Arrest, Arraignment, Trial, Sentencing, and so on. Each of these frames is related to the Criminal_process frame via the SubFrame relation in the database. Moreover, subframes (of the same complex frame) are related to each other through their ordering. (See Narayanan et al. (2002) and Scheffczyk et al. (2006).) Note that the FEs of the subframes need not map to those of other subframes. So, in this same group of frames, the Arrest frame includes an ARRESTING_AUTHORITY which is not included in the Trial frame.

Notice that a given subframe may itself be a complex frame. For example, the Trial frame is a subframe of the Criminal_process frame, and has its own rich structure, some of which can be decomposed into simpler frames that are related to each other. A Trial is made up of court appearances, and involves opening arguments, presentation of evidence and testimony, and closing arguments. The system of subframe links is also quite complex. At present, the subframe relation is not indicated in every relevant case.

6.1.4 Precedes

This relation occurs only between two Component frames of a single Complex frame, i.e. as extra information associated with a set of Subframe relations. It specifies the sequence of states and events that are definitional for a certain state-of-affairs. Most Subframe relations will naturally have precedence relations, as can be seen in the foregoing diagram of the subframes of the Criminal_process frame (6.3), in which the Precedes relations are indicated by the black lateral arrows.

This is the only frame relation for which cycles are allowed. For example, in the subframes of the Sleep_wake_cycle frame (shown in Figure 6.4), Being_awake precedes Fall_asleep, which precedes Sleep, which precedes Waking_up or Getting_up, which in turn precedes the first frame, Being_awake.

6.1.5 Causative_of and Inchoative_of

We record the especially close and fairly systematic non-inheritance relationships between stative frames and the inchoative and causative frames which refer to them using the frame-to-frame relations **Causative_of** and **Inchoative_of**. Consider the following frames: Position_on_a_scale, Change_position_on_a_scale, and

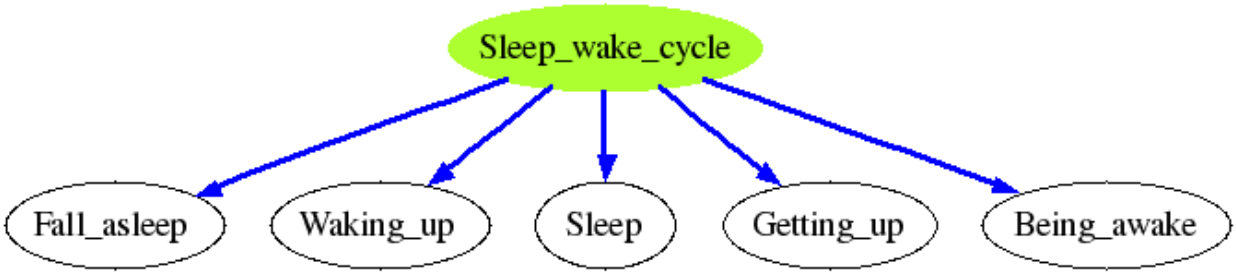

Figure 6.4: Subframes of the Sleep_wake_cycle frame

Cause_change_of_scalar_position. The following sentences illustrate the kind distinction captured in these frames:

(1) [Cause_change_of_scalar_position] (raise.v) Billie Blount **raised** taxes on farmers 18 times in 2002!

(2) [Change_position_on_a_scale] (rise.v) During the Elizabethan age, there was an increased emphasis on genealogy in the heralds' work as the gentry class **rose** in importance.

(3) [Position_on_a_scale] (high.a) Most fish from lakes is too **high** in mercury.

Frames which participate in these relations as Causatives should inherit from the Transitive_action frame, Inchoatives should inherit from Event, and States from State or Gradable_attribute.[6] Many of these inheritance relations have not yet been made.

6.1.6 Using

Often a particular frame makes reference in a very general kind of way to the structure of a more abstract, schematic frame. Since the creation of the more specific Perspective_on relation (see 6.1.2), the Using relation is used almost exclusively for cases in which a part of the scene evoked by the Child refers to the Parent frame. For example, Volubility uses the Communication frame, since Volubility describes a quantification of communication events.

It is possible for a frame to use more than one frame. An example of this situation is the Judgment_communication frame, shown in Figure 6.5 below. It uses both the Judgment frame and the Statement frame. The Judgment_communication frame does not inherit Judgment because it is not a simple subtype of a purely cognitive state. Judgment_communication does not inherit Statement either since it distributes the content of the Statement frame's MESSAGE frame element over two frame elements, EVALUEE and REASON.

6.1.7 See_also

In cases where there are groups of frames which are similar and should be carefully differentiated, compared, and contrasted, each of the frames in question has a SeeAlso relation with a representative member of the group. In the frame definition of the representative member, there will be a comparison which will contrast the frames to make clear the intended boundaries between them. For example, since the Scrutiny and Seeking frames are similar, there is a SeeAlso relation from Scrutiny to Seeking, and text in the Seeking frame that explains the difference.

This relationship does not imply any particular relationship between the Frame Elements or subframes of the frames involved.

[6]For many frames inheriting from Gradable_attribute, Inchoative or Causative frames corresponding to them should exist but have not yet been created. E.g. for the Age frame, which has LUs like *old.a*, there should be corresponding frames for the verb *age*, namely a Causative frame we might call *Cause_change_of_age and an Inchoative frame *Change_of_age.

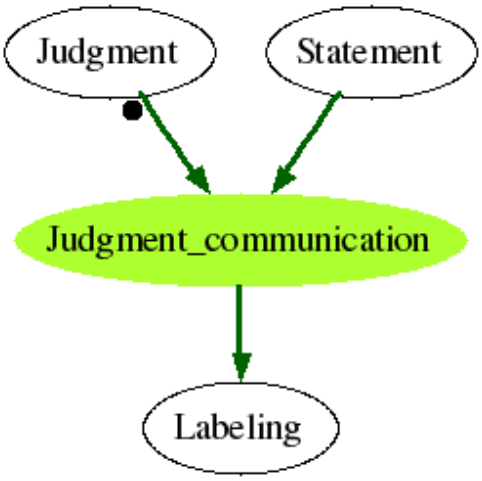

Figure 6.5: Judgment_communication's Using relations

6.2 Semantic Type

The general use of semantic types in the FrameNet project is to record information that is not representable in our frame and frame element hierarchies. In this section there is a detailed description of each major category of the semantic type hierarchy, which is broadly split by function. In specific, the functions we currently employ semantic types for are:

- Indicating the basic typing of fillers of frame elements, e.g. "Sentient" for the COGNIZER FE. These basic types are constrained by the frame hierarchy (see section 6.3.5.1), but not predictable from it since frame elements which are arbitrarily far away according to the frame hierarchy, such as the EXPERIENCER of Perception_body and the PERPETRATOR of the Piracy frame, are often marked as the same semantic type (in this case, Sentient). This kind of semantic type is designed primarily to aid frame parsing and automatic FE recognition.

- Useful, functional marking on frames, such as the type "Non-lexical" on frames which are present purely to participate in Inheritance, Subframe, or Using relations with other frames. This kind of property is actually a meta-description, not a fact about the semantics of the frame at all, and thus independent of the hierarchy. In fact, a frame which in English is Non-lexical might well have associated LUs in another language.

- Marking important dimensions of semantic variation among the lexical units in a frame that are not related to the kind of semantic combinatorial possibilities that we use for making frame distinctions (see Chapter 2). For instance, in the Judgment frame the difference between LUs such as *praise.v* and *criticize.v* in terms of the negative versus positive evaluation of the EVALUEE is marked with the semantic types Positive_judgment and Negative_judgment, respectively.

The most interesting function of semantic types for human users is the third one, recording important semantic differences between lexical units that recur within several frames.

For example, "Positive_judgement" and "Negative_judgement" semantic types, indicating the speaker's attitude toward a situation, can be applied to lexical units across a range of frames. (Note that the term *speaker* may either refer to a frame element such as SPEAKER or COGNIZER, or to the author of the utterance containing the lexical unit.) We capture the distinction by marking LUs like the aforementioned Judgment verb *praise*, the Experiencer_subj verb *like*, and the Frugality adjective *generous* as "Positive_judgement" and, by contrast, marking "Negative_judgement" on some other words from the same frames, e.g. *criticize* and *hate*, and *stingy*.

For more description and further examples, see the following subsections.

6.2.1 Ontological Types

These semantic types are used to classify the denotation of lexical units, frames, and frame elements. On *lexical units*, they classify the denotation of LUs, usually cross-cutting their classification by frames.[7] When, on the other hand, such a distinction corresponds to a frame distinction, the *frame* is labeled with the type, which then signifies that every LU of the frame could be labeled with an equally or more specific Type. For *FEs*, semantic types actually categorize the sort of *filler* that is expected in the FE, rather than classifying what kind of role the FE is.[8]

Most of these types correspond directly to synset nodes of WordNet, and can be mapped onto ontologies, e.g. Cyc. The types are related to each other via subtype relations, which are logically equivalent to Inheritance or is-a. However, the FrameNet semantic type hierarchy is not guaranteed (or even likely) to correspond exactly to that of WordNet, Cyc, or any other resource.

One example of a semantic type which has been used for frames, FEs, and LUs is *Container*. This corresponds most closely to the WordNet synset node *container* and the Cyc Container Collection node.[9] The Container frame, unsurprisingly, contains LUs which refer to containers like jar.n, box.n, etc., and thus both the frame and the LUs in it can be labeled with the semantic type *Container*. The SOURCE of the Emptying frame is also labeled with the ST *Container*. This means that head-nouns of this FE are likely to be the sorts of words that belong as LUs in the Container frame. Pronouns or other kinds of nouns may also occur as fillers of the FE, in which case they are construed as containers, as in the following:

(4) She **emptied** *the lid* back into the bottle.

In a neutral context, we wouldn't expect a lid to itself be a container. Not every lid is at all readily describable as a *Container*, e.g. the lid of a tin can. Regardless of whether a filler of this role is inherently a *Container*, its occurrence as a filler of the role means that it is being *used* as a *Container* in the current context, and has the properties necessary for that.[10]

6.2.2 Framal_type

These types, as mentioned above, can only be applied to frames and are **not** to be interpreted as applying to LUs of a frame or any connected frames, including Inheritance daughters (which normally should match any semantic types that are placed on their parent frames).

6.2.2.1 Non-lexical_frame

Such frames have no lexical units and are present purely to connect two (or more) frames semantically. One example is the Post_getting frame, connected to Getting via a Precedes relation and connected to Possession via an Inheritance relation. This allows us to succinctly encode the fact that the state following "getting X" (the Getting frame) is "having X" (the Possession frame).

6.2.2.2 Non-perspectivalized_frame

This semantic type is used for frames which have a great diversity of lexical units, all of which share a kind of scene as a background. Such frames do not have a consistent set of FEs for the targets, a consistent time assigned to the events or participants, or (most especially) a consistent point-of-view between targets. An example of this type of frame is the Performers_and_roles frame, which contains such diverse LUs as *co-star.v*, *feature.v*, and *as.prep*. Like the Biframal_LU types, this semantic type is intended as a time-saving measure. All such frames could be split up into smaller frames with a consistent perspective, but these frames would contain very few LUs. (See 6.2.3.3 on Biframal_LUs.)

[7] For example, the semantic type Body_of_water is applied to some LUs in the Biological_area frame, e.g. *bog.n* and to LUs in the Natural_features frame, e.g. *bay.n*.

[8] We have implemented the typing on roles themselves (such as whether they are complements or adjuncts) using the separate Coreness status feature (see Section 3.2.1).

[9] These are not quite equivalent to the FN definition. Cyc, for example defines Container Collection as referring to "tangible objects whose primary function is to contain something" (`http://opencyc1.cyc.com:3602/cgi-bin/cyccgi/cg?cb-start`) rather than the broader FrameNet definition which allows any entity which is construed as containing something.

[10] In some cases, a filler occurs in a particular role despite the fact that it has features which directly contradict the semantic type assigned to a role. For example, in "She filled her days with meaningless tasks", *day.n* is a non-physical entity and thus clashes with the type *Container* which inherits from *Physical_entity*. All such cases will need to be interpreted metaphorically or metonymically by resources beyond the scope of FrameNet.

6.2.3 Lexical_type

There are a number of labels applied to LUs or frames which do not refer directly to the kind of entity denoted by an LU or the semantics of a frame. These demand, rather, some more complicated interpretation of the LU or frame that they are attached to. How entities with such types are to be interpreted is different for each of the types that are described in this section (and in the following section on Framal types), and in some cases the label even indicates that an LU should be interpreted in a different frame from the one that it is attached to (see Guest_LU types below 6.2.3.3), so LUs with these types, if interpreted as subtypes of the semantics of the frame according to the normal FrameNet logic (see 6.3.4.1), can be very misleading. (This means that for some purposes, it would be more sensible to leave out items with these types attached than to include them without further comment.) A description of each of the types follows.

6.2.3.1 Transparent_nouns

LUs marked as **transparent nouns** have an unusual sort of semantics since, unlike most nouns, their primary function is to give some kind of description of another noun, as seen in the examples with *kind*, *number*, and *side* below:

(5) I **read** that kind of book in college.
(6) Monotremes **lay** a number of eggs per clutch.
(7) Put it **on** the left side of the shelf.

Specifically, transparent nouns describe either 1.) information which is not typically inherent in the nouns they describe, such as quantity, grouping, and shape, or 2.) information which can be left implicit in the interpretation of nouns when they denote an instance of an entity, such as a.) instancing itself (e.g. "this bat" = "an instance of a bat"), b.) subtyping (e.g. "this bat" = "a type of bat", since an instance is a maximally specific type, and thus automatically a subtype of its class, or c.) picking out subparts (e.g. "with a knife" = "with the edge/point of a knife"), as an **active zone** subpart of an entity is often meant even when the whole entity is mentioned.

The descriptive aspects of meaning profiled by transparent nouns are usually of little consequence for determining acceptability as fillers of predicate roles; rather more important is the semantics of the entity being described. This can be seen in the examples below, in which the above sentences are summarized/paraphrased by omitting the transparent noun, leaving only the described entity:

(8) I **read** those books in college.
(9) Monotremes **lay** many eggs per clutch.
(10) Put it **on** the shelf, towards the left.

We judge that recognition of these LUs is vital for correct cataloging of FE fillers, summarization, paraphrase, etc. since they violate the normal rule that the syntactic head of a phrase is the semantic head. For most purposes, transparent nouns can simply be omitted from the analysis of a sentence.

However, note that transparent nouns are not always transparent. In the right semantic contexts, they are so to speak "opaque" and dominate the frame of their dependent (as most targets do); in such cases the meaning of the so-called transparent noun itself rather than that of its dependent is selected by a predicator. In particular, this is the case with governing predicates that evoke an open proposition or question, e.g. determine, find out, measure, etc.

(11) Archaeologists have been determining the **number/shape/part/type** of tools used by the Maya.

Here none of the transparent nouns can be used with quite the same meaning as the that of the described entity alone:

(12) Archaeologists have been determining the tools used by the Maya.

In addition, transparent nouns convey important, if optional, information like quantity and configuration. For tasks concerned with these dimensions of meaning, transparent nouns should be treated like any other target.

6.2.3.2 Bound_LU

Some LUs of a frame cannot be used unless they co-occur with the right other words. The limiting case, where a precise group of words must all co-occur in order to evoke a frame, is handled in FrameNet by the use of multi-word lexical units (3.1). But there are many cases where an LU may evoke the frame when occurring with any of a number of different words, or even when occurring with a class of words. Such LUs are to be marked with the *Bound_LU* semantic type.

There are two subtypes of Bound LU that we recognize in FrameNet: Supports and Bound_dependents.

Support This subtype of bound lexical unit is applied to LUs that only evoke a frame when they are used as Supports of certain dependents, as exemplified by *give.v* in sentences like the following:

(13) Receiving the notification so late almost **gave** me a heart attack.

In this sentence, *give* is used approximately synonymously with *cause*:

(14) Receiving the notification so late almost **caused** (me to have) a heart attack.

This use of *give*, then, should go in the Causation frame, but it is readily apparent that this meaning is not possible in many (if not most) contexts:

(15) ??Releasing these old files nearly **gave** a disaster.
(16) Releasing these old files nearly **caused** a disaster.
(17) *What **gave** these events?
(18) What **caused** these events?

As implied above, *give* occurs in this meaning only when it is to be annotated as a Support Verb (see 3.2.7.1). (Note that there are other support uses of *give* which are not equivalent to *cause*, e.g. *give a laugh*, *give a speech*.)

In principle, though not yet in fact, all supports with any semantics not included in the target noun, adjective, or preposition that they are associated with, such as causative supports, inchoative supports, and point-of-view-shift supports, should be annotated separately in a causative, inchoative, or point-of-view specifying frame.

Bound_dependent In addition to fixed expressions and support constructions, there is a small class of other LUs which only occur as dependents of a limited set of governors. This semantic type has not been included on LUs in the current data release, but it is intended to cover semi-productive LUs such as *attention.n* in the Posture frame (in phrases like *(stand) at/to attention*), *bind.n* in the Predicament frame (in phrases like *in/out of a bind*), and idiosyncratic degree modifiers from the yet-to-be-made *Degree frame (for LUs like *very.adv*), e.g. *dirt* in *dirt poor*, *crystal* in *crystal clear*, *sopping* in *sopping wet*, etc. These LUs can be added to appropriate frames and given a Bound_dependent type, from which the user should infer that they are only usable when they have particular governors. The appropriate governors must be gleaned or generalized from the usage in the annotation data.

6.2.3.3 Biframal_LU

LUs of a frame normally have a semantics which is a subtype of the semantics of the frame. However, when LUs are marked with a Biframal_LU type, they denote something related, but not equal to the semantics of the frame. In all cases where an LU is marked as biframal, we could have made a separate frame to more narrowly characterize the meaning of the LU. These types have been defined as a time and resource-saving measure, to avoid having to create these separate frames for tiny groups of LUs; for example, *(un)intelligible.a*, despite the close relationship with the Grasp frame, would otherwise require a separate frame.

Another way of looking at this is that the normal relation of the semantics of the LU to the semantics of the frame is Inheritance, while the relation of a biframal LU's semantics to the frame's semantics is generally Using. Except for the Guest_LU type, each of the following subtypes define what frame the biframal LU inherits from. (See below for Guest_LU.)

We tend to avoid using these semantic types, since they complicate the interpretation of our data, and in many cases there is a problematic mismatch in the FEs allowed for the biframal LU and the other LUs of the frame.[11]

[11] E.g., *gunman.n* is included as an Agentive_noun LU in the Bearing_arms frame, but this frame does not include all of the

Guest_LU This type is applied to LUs which are only tangentially used in the host frame, and whose interpretation is still largely dependent on their membership in some other frame, as demonstrated in Example 6.2.3.3.

(19) Statement: " You'll never catch up," she **grinned**.
(20) Gesture: She **nodded** him through the door .
(21) Caused_motion: He **sneezed** the handkerchief off the table.

Thus, for example, *grin.v* can be included in the Statement frame, but any full understanding of this sentence requires an understanding of the Making_faces frame.

Agentive_noun This type is used to mark LUs which denote the AGENT[12] of the frame in which they occur, as in *murderer.n* in the Killing frame. They can be thought of as virtually inheriting from People and using the frame they are listed in; the PERSON FE of the People frame is bound to the same FE that is bound to the AGENT FE of the host frame.

Participating_entity This type marks LUs which denote the PATIENT[13] of the frame in which they occur, e.g. *possession_((entity)).n* in the Possession frame, *victim.n* in the Crime frame, and *knowledge.n* in the Awareness frame. These LUs can be thought of as inheriting from the Entity frame, and as using the host frame, with the PATIENT (or equivalent) FE of the host frame bound to the same FE as is bound to the ENTITY FE of the Entity frame.

Tendency_grading_LU [14]

This type marks LUs which denote the propensity (of something) to be a PATIENT (or similar FE, see 6.2.3.3) in the host frame; e.g. *verifiable* in the Verification frame is marked as a tendency-grading LU since it concerns something's tendency/ability to be verified. The semantics of a tendency-grading LU can be understood as a virtual frame inheriting from Inclination and using the host frame.

The most common subtype of these LUs are abilitative passives, usually constructed morphologically from a verb (X) + *(a)ble*, which denote the propensity of an entity to be Xed. Normal, productively formed examples include *findable* (Locating), *usable* (Using), *likable* (Experiencer_subj), and (with negative prefix) *unfixable* (Resolve_problem). There are also semantically indistinguishable examples which are not productively formed from verbs, e.g. *intelligible* (Grasp). These LUs are listed in frames that inherit from the causal Transitive_action frame.

In addition, there are some LUs which do not directly make reference to a causal frame, but rather refer to an inchoative frame, e.g. *fragile.a* (Fragment), and *mortal.a* (Death).

6.2.3.4 Affect_describing: Positive- and Negative_judgement

Many LUs describe the positive or negative opinion of an implicit or explicit JUDGE, without actually necessarily being of similar type otherwise, e.g. *like.v* in the Experiencer_subj frame, and *stingy.a* in the Stinginess frame. This type labels such LUs so that they can be compared despite their dissimilarities. In addition to the positive or negative dimension, there are two main types, distinguished by *who* is identified as having a high or low regard for something: 1.) Emotion-related targets, where the EXPERIENCER (or descendant of EXPERIENCER) is explicitly identified as the opinion-holder and the opinion concerns an overt STIMULUS or EVALUEE, e.g. *approve.v*, and 2.) other targets where the user of the word in the discourse is the opinion-holder and the opinion concerns the target predication as a whole, e.g. *thrifty.a*. All LUs in frames which inherit from or use the Emotions frame, especially the Judgment family of frames, fit in type

FEs appropriate for annotating people, like ETHNICITY, so that "Somali" in "Somali gunmen" is not annotatable; this problem is even more obvious with a word like *breakable.a* in the Render_non-functional frame, since it requires a DEGREE FE which is not available for other members of the frame. In such cases, we have the unenviable decision between adding a frame element which is not useful for most targets of the frame, and ignoring an important FE; in practice, we follow the latter approach and do not include FEs which are only possible for biframal LUs.

[12]For the purposes of this section, AGENT is taken to refer to any frame element which is connected to an AGENT FE via inheritance, even if this daughter is not called an AGENT; e.g. the COOK Agentive FE in the Apply_heat frame. Note that if an LU simultaneously denotes an FE inheriting from AGENT and any other FE (e.g. *braggart.n* in the Bragging frame, which also denotes the EVALUEE of Judgment) then Agentive_noun is used.

[13]For the purposes of this section, PATIENT is taken to refer to any frame element which is connected to a PATIENT FE via inheritance, even if this daughter is not called a PATIENT.

[14]This type is renamed and broadened from the Abilitative_passive type in previous releases.

1, with the COGNIZER, SPEAKER, or EXPERIENCER as the opinion-holder, whereas all other targets express the opinion of the individual using the word.[15]

In words of the second type, where it is the language-user whose view is expressed, the full semantics of the LU could be thought of in terms of a virtual frame which inherits from the host frame and uses the Regard frame, with the COGNIZER specified as the language-user, the EVALUEE identified as the whole predicate, and the JUDGMENT identified as positive for Positive_judgement and negative for Negative_judgement. Thus the target *bony.a* of the Body_description_holistic frame in 6.2.3.4 below indicates not only that someone was thin (which is at least partly specified by the meaning of the frame) but also the person declaring this state considers this undesirable.[16]

(22) Ignacio said she was looking **bony**.

6.2.3.5 LU_with_FE_specified

These types are employed on LUs to indicate that the LU inherently specifies some information about an FE. Although the types of information that LUs can specify about FEs is considerably broader (including most saliently filler-types for an FE which are particular to an LU, as in *tie.v* in the Attaching frame), we have so far used only two basic kinds, which specify DEGREE and sensory modality respectively.

Degree_specified_LU (Negative and End_of_scale) These types are used for LUs whose framal semantics provides a DEGREE FE. All such LUs inherently describe a particular ATTRIBUTE of an ENTITY as deviating in a particular direction from the norm. (This norm is usually the expected value of the ATTRIBUTE for the type of ENTITY described, but may be a norm for some more abstract supertype of the ENTITY. See the Gradable_attributes frame.) LUs should pre-specify the direction and difference from the norm; Negative indicates the direction and End_of_scale indicates maximal difference.

LUs which have the type Negative specify that the direction of deviation is in the negative direction, i.e. either less in quantity or desirability (or at any rate in the opposite direction from the unmarked description given in the frame definition). These LUs thus form antonyms to unmarked LUs in the same frames. Some select cases are presented in Table 6.2. Note that many LUs that should receive this type have not been marked in the FrameNet data.

Frame	Negative LU	Positive LU
Ambient_temperature	cold.a	hot.a
Experiencer_subject	hate.v	love.v
Mental_property	stupid.a	smart.a

Table 6.2: Antonymy in select frames

End_of_scale marks LUs that characterize an ENTITY as having a maximal value for an ATTRIBUTE, e.g. *fabulous* in the Desirability frame. Such LUs are not usable with normal DEGREE modifiers like *very*.

Sensory_related_LU LUs marked with these types inherently reference some type of sensory experience, either directly denoting such an experience, e.g. *see.v*, *look.v*, or other LUs in the Perception_experience or Perception_active frames, or they use such a frame, and thus imply the occurrence of a perceptual event, e.g. *tasty*. In any case, the important function of the subtypes of this type (viz. Visual_modality, Tactile_modality, etc.) is to identify the subtype of sensory experience which is inherent to an LU. Some examples of the use of the Visual_modality type:

(23) *see.v* in Perception_experience
(24) *glance.v* in Perception_active
(25) *shiny.a* in Location_of_light
(26) *ugly.a* in Aesthetics

[15] In some few cases, such as *fob/pawn off.v*, it seems that the negative judgement (of the language-user on the THEME in this case) is also ascribed to another participant in the frame (in this case the DONOR); this fact is not currently representable in the FrameNet data.

[16] In this example, since the word is embedded as an FE in the context of a speech verb, a full analysis might conclude that the explicit SPEAKER is also the opinion-holder, but this type of inference lies strictly in the domain of mental-spaces analysis rather and outside frame analysis proper.

6.3 How to use relations

This section outlines some of the major uses for the relations and semantic types. Section 6.3.1 is primarily devoted to describing the utility of the relations for humans using FrameNet as a reference, Section 6.3.2 is of use to both humans and automatic programs, whereas the other sections are of more interest to developers intending to use FrameNet for computational purposes. Virtually any computational use of the FrameNet relation information relies on and presupposes semantic parsing of texts–a process not discussed here, but well covered in several publications (e.g., Erk and Padó (2006) and Litkowski (2004)).

6.3.1 Looking up related words

One obvious benefit of defining relations between frames is that it enables the look-up of words that are related. For example, we can look at the word *murder.v* which occurs in the Killing frame. Obviously, there's some level of similarity with the other members of the frame, since they all involve something causing an entity to die. This sort of relationship might be captured in a good thesaurus under an entry for *kill.v*. However, if we follow the Causative_of relation from the Killing frame to the Death frame, we find a whole new group of words, also about death, e.g. *die* and *pass away*, but in this case focusing on the change of state of the PROTAGONIST. This kind of relation is neither synonymy nor antonymy, and therefore not to be found in a thesaurus.

6.3.2 Comparing to other systems of semantic annotation

In Fillmore's earlier work Fillmore (1968, 1977), a case was made for the universality of certain types of semantic roles, a concept which was further developed and is now enshrined as the theta role system of many syntactic/semantic formalisms. These roles include such labels as Agent, Instrumental, and Objective (roughly corresponding to Agent, Instrument, and Patient in other formulations). However, as the description of the semantics of Lexical Units has progressed, it has become apparent that the theta-role and original case-role account covers only a subset of the full set of roles.

We now take it that theta roles should be mapped to FEs in high-level, abstract frames like Transitive_action, which has FEs like AGENT and PATIENT. The relevance or irrelevance that these labels have for the roles of more specific predicates like *break.v* (in the Cause_to_fragment frame) or *resemble.v* (in the Similarity frame) is explained explicitly by the inheritance or non-inheritance of the AGENT and PATIENT frame elements in the relevant frames.

There are inherent problems to reducing our role-set of Frame Elements to the considerably smaller (and inarguably more computationally tractable) set of theta roles often used. One of these is deciding on the initial set of theta roles – a well-known problem in the theta-role literature. The Frame Elements that we define, however, are more immediately verifiable. Presumably, then, any theta-role system proposed to cover all predicates should allow us to specify, in a simple way, which FEs should be mapped to which theta-roles. Of the theta-role systems known to us, none allows any simple mapping to high-level FEs (and thence to the FEs that inherit from them) without covering some FEs multiple times and/or leaving some FEs uncovered, unless there is an unsatisfactory catch-all theta-role.

6.3.3 Paraphrase and translation

In many ways, paraphrasing is at the core of what we intend FrameNet to facilitate. A properly powerful ability to paraphrase enables many of the other goals of semantic NLP, including Question Answering, Summarization, and Translation. Question Answering can be thought of as looking in a corpus to find a paraphrase, but with real information filled in for the questioned FE. Summarization is equivalent to paraphrase of a text, but with the strategic omission of information from FEs and targets. Translation is paraphrasing with the limitation that all the resulting paraphrase must be in the target language.[17]

One of the basic insights behind FrameNet is that grouping words according to the scenes that they evoke, regardless of whether they are synonyms, antonyms, or some other relation to each other, groups words that are useful for paraphrasing. In particular, since FrameNet lists words together despite part-of-speech differences (unlike WordNet), paraphrases involving an interchange of noun, verb, adjective, or preposition are (in principle) discoverable with the FrameNet data. (See Sect. 6.3.3.1.)

[17]This requires FrameNet-style data for both source and target language; this currently limits such efforts to English, German, Japanese, and Spanish.

6.3.3.1 Differences in LUs: Semantic Types and WN

Despite the fact that FrameNet is built to facilitate paraphrase, more specific relationships between the words usually need to be ascertained to recognize or generate actual instances of paraphrases. What kind of paraphrase results is dependent on whether a proposed paraphrasing target is an antonym, exact synonym, subtype, or supertype of the initial target to be paraphrased.[18] We have some of this more specific information (see the Semantic type "Negative", 6.2.3.5), but for the most part, we have envisioned the information that we record as complementary to WordNet.

As WordNet has a reasonable coverage of synonymy, antonymy, and subtyping ("is-a") relations between words, we have tried to avoid reinventing the wheel by duplicating this information in FrameNet. However, for tasks that require information from both sources, integrating the information would depend on mapping FrameNet LUs to words in WordNet synsets in some way. This has not been consistently accomplished anywhere, although work has been done by several researchers.[19]

Once a proposed list of paraphrases for a starting LU is established (however buggy), then targets can be paraphrased with other words which are synonyms, supertypes, subtypes, or (with the addition of negation) antonyms in the same frame; this excludes, then, targets which are "sisters" or "cousins", i.e. which are subtypes of some supertype of the original target.

Working without a WN-FN mapping, antonyms can at least provisionally be established if some members of the frame have the Semantic type Negative. Such targets are antonyms of unmarked targets in the same frame. In addition, targets with FE incorporation should be considered subtypes of targets in the frame without incorporated FEs. In the absence of any other information, all other targets could be considered synonyms, although this will be erroneous in many cases.

Degree differences, especially antonymy Paraphrasing an initial target with a target that is an antonym requires adding in negation, as can be seen in 27 and 28 with antonymic targets from the Possession frame:

(27) Initial sentence: She may **lack** the money for a more nutritious meal.

(28) Paraphrase: She may *not* **have** the money for a more nutritious meal.

Supertypes and synonyms When paraphrasing a more specific initial target with a more general target from the frame, no adjustment is generally necessary, just as with a synonym, as seen in examples 29 and 30 from the Self_motion frame:

(29) Initial sentence: I enjoyed watching the giraffes **saunter** by.

(30) Paraphrase: I enjoyed watching the giraffes **walk** by.

Supertype paraphrase works across frames as well; compare 31 from the Motion frame, which Self_motion inherits from:

(31) Paraphrase: I enjoyed watching the giraffes **go** by.

Interchanging part-of-speech of targets The simplest type of interchange of part-of-speech of targets does not require any change in the embedding context. This is most often the case when a target combines with a support to "simulate" a different part-of-speech. For example, a noun + support verb can often paraphrase a verb target (example from the Bragging frame):

[18] Some pairs of targets within a frame will not be fit paraphrases at all. This occurs when the targets have differing semantic relations (synonymy, antonymy, hyponymy, etc.) to a more generic concept within the frame; for example, the Subject_stimulus frame has such relative incomparables as *funny.a*, *shocking.a*, *sad.a*, *encouraging*, etc. These have no sensible paraphrase relationship with each other, except for the fact that they are all subtypes of emotional description focusing on the EXPERIENCER. Something similar can be said for frames like Biological_area with words like *forest.n* and *bog.n*.

[19] Martha Palmer and Andy Dolbey are working on the most thoroughgoing effort to make a broad alignment of WordNet, VerbNet, and FrameNet (Martha Palmer, personal communication), but this project has not yet produced publications.

Mapping in the other direction, Aljoscha Burchardt and others have created the "WN detour to FrameNet", which connects WN synsets to FN frames (`http://www.coli.uni-saarland.de/albu/papers/gnws05_burchardt_erk_frank-final.pdf` and `http://www.coli.uni-saarland.de/albu/cgi-bin/FN-Detour.cgi`).

Several other efforts have used a knowledge-engineering approach to put FN data together with WN (Shi and Mihalcea (2005)) or VerbNet and PropBank (Giuglea and Moschitti (2004)).

(32) Initial sentence: Eustace **boasted** uncomfortably.
(33) Paraphrase: Eustace *made* an uncomfortable **boast**.

In 32 and 33, *boast.v* and Support Verb + *boast.n* equivalently fill the finite-verb slot of the sentence.

They may also equivalently fill exactly the same valence pattern of a governing verb, as seen in 34 and 35. In this case, both noun and verb fill the VP-to realization of the GOAL FE of the verb *try.v* in the Attempt frame:

(34) Initial sentence: She *tried* to **boast** about her years of teaching experience.
(35) Paraphrase: She *tried* to make some **boast** about her years of teaching experience.

A noun + support preposition can paraphrase an adjective or adverb target:

(36) Initial sentence: He sat *in* some **discomfiture** as she explained.
(37) Paraphrase: He sat somewhat **discomfited** as she explained.

Other kinds of interchange of noun, adjective, verb, or preposition require changing the Phrase Type of the phrase that they occur in. This is only possible when the governing context has two different valence patterns with different phrase types for the FE filled by the target (see 6.3.3.2):

(38) Initial sentence: We all **want** to **succeed**.
(39) Paraphrase: We all **want success**.

Here, the paraphrasing of *succeed.v* with *success.n* is only possible because the EVENT FE of the target *want.v* in the Desiring frame can be realized as either an infinitival complement or an NP object.

6.3.3.2 Paraphrasing FEs with different PTs

One strategy of paraphrase is to exchange the valence pattern for the realization of FEs for a target; a typical case is the Dative Alternation:

(40) Initial sentence: **Hand** me that spatula.
(41) Paraphrase: **Hand** that spatula to me.

In this example (from the Giving frame), the RECIPIENT FE is filled either with an NP or a PP(to). Since the alternation of NP and PP merely involves the presence or absence of the marker (i.e. *to*), it is easily possible to paraphrase these mechanically.

In principle, for most possible pairings of PTs, there are cases when a particular FE will vary between the two PTs across a frame. All such cases provide potential paraphrases across phrase types.

For example, the Phrase Type "PP-ing(about)" in the sentence "She thought about returning" can be paraphrased in the sentence "She considered returning" with the PT "VPing", or in the sentence "She considered a return" with the PT "NP", or in "Her contemplation of a return" as "PP(of)", etc. These interchanges can be selected from the phrase types listed for a particular FE realization in the valence tables of the lexical entries. Thus, this type of paraphrase generation is relatively independent of the target choice (discussed in 6.3.3.1), limited only after the fact by what phrase types are available for the chosen target.

Many of the interchanges cannot be achieved by simply varying a marker, e.g. changing an NP to a VP or vice-versa. This more complicated type of interchange could only be accomplished by paraphrasing the targets inside the phrase, e.g. by paraphrasing the head noun of the NP with a verb from the same frame to make a VP.

The straightforward interchanges of PT can be produced mechanically by adding, subtracting, or changing a marking word (like a preposition) or changing the morphology of the head of the phrase (e.g. from present participle to infinitive). The paraphrasability of the unparenthesized PTs below in Figure 6.6 is merely a matter of changing the morphological marking on the head and/or changing the syntactic markers (such as prepositions, quotation marks, or the infinitive-marker to). For many other patterns (indicated in parenthesis below), there are very common constructions (such as copularization of adjectives) which allow an interchange of PTs.

- NP Poss PP (N) (A)
- VPing VPto VPbrst VPed VPfin PPing (Swhether Sinterrog PPinterrog)
- Sinterrog Swhether PPinterrog QUO
- Sing Sto Sforto Sbrst Sfin QUO
- AJP (AVP Sub) (Srel VPtorel VPed)
- QUANT (Srel)
- Sabs (Sub)

Figure 6.6: Simple phrase-type valence paraphrases

6.3.3.3 Paraphrasing by omission

Many FEs are optionally syntactically expressed, and many (less informative) paraphrases which omit them can be easily generated.

All peripheral FEs and non-target incorporated FEs are freely omissible, thus (omitting the TIME FE of the Giving frame):

(42) Initial sentence: **Hand** me that spatula now.
(43) Paraphrase: **Hand** me that spatula.

In addition, FEs that ever occur with an LU as INI should be considered omissible.[20]

FEs that occur with an LU as DNI are omissible under essentially the same conditions as would license an anaphoric pronoun, deictic adverb, or the like for the same FE.

FEs that occur with CNI omission are omissible only in certain constructional contexts. The FrameNet data currently do very little to explicate what these contexts are, although they are generally well known syntactic phenomena.[21]

6.3.4 Inferencing

We intend that the Frame Relations, Frame Element mappings, and Lexical Units provide a significant amount of information for doing inferencing on natural language. We do not provide every kind of information that an inferencing engine might require (especially any kind of quantitative information), but the frame-to-frame relations should provide a good basis for further specifications (using X-Schemas, predicate logic, etc.) which are useful to existing systems, especially in combination with information from WordNet. To this end, an OWL version of these relations is included in the data release.

This section describes some of the kinds of inferences that are intended by the definitions we give to the various relations.[22]

The most basic summarization of the logic of FrameNet is that Frames describe classes of situations, the semantics of LUs are subclasses of the Frames, and non-Extrathematic FEs are classes that are arguments of the Frame classes.[23] An annotation set for a sentence generally describes an instance of the subclass associated with an LU as well as instances of each of its associated FE classes.

Inheritance has a special place among the frame relations. It is to be interpreted as a subtype or "is-a" relation between frame classes. This constrains the Child frame to have all of the properties of the Parent frame.

[20] In many cases, we have annotated the omitted FEs in generic sentences with INI. Such omissions would be more properly labeled CNI, on which see below.

[21] Correctly using CNI information for paraphrase would require specific methods for handling the constructions commonly referenced in the FrameNet data, i.e. Passives, Imperatives, Instructional imperatives, and Generics (including the common omission of external arguments with infinitives and participles).

[22] Note that the See_also relation has no formal semantic definition whatsoever, and thus is not intended for machine processing. It will not be mentioned further in this section.

[23] As mentioned in section 6.3.4.1, Extra-thematic FEs effectively evoke their own frame, separate from the frame of a target LU. The Extra-thematic FE itself is a Frame Element of this Extra-thematic frame. Other frame elements of the Extra-thematic frame should correspond to some subset of the Frame Elements of the target LU's frame, but exactly which subset of FEs is not identified in the current data release.

Although it is difficult to generalize across all of the relations, it is usually true that an instance of a Sub frame implies the existence of an instance of a Super frame. In the case of a Theft, which inherits from Committing_crime, an instance of Theft is obviously an instance of Committing_crime. For the Precedes relation, however, this is not exactly true .[24] For the Subframe relation, the existence of a Sub frame instance does imply the relevance of a Complex frame instance, but the reverse is not true. So, for example, if there is an instance of the Sentencing frame, then we can instantiate an instance of the Criminal_process frame. We can then also instantiate all of the subframes of Criminal_process which must precede Sentencing, including Trial, Arraignment, etc., but we cannot instantiate the Appeal subframe of Criminal_process which follows Sentencing.

6.3.4.1 Logical relations of FEs and LUs to Frames

The relations of FEs and LUs to frames are intertwined with our Frame relations in various ways. Some clarification and formalization of these concepts is necessary to simplify their logical description:

- The term "Frame Element" has two meanings: the relation itself, and the filler of the relation. When we describe the Coreness status of an FE (see Sect. 3.2.1), we are describing the relation; when we describe the Ontological type on an FE (see Sect. 6.2.1) we mean the type of the filler. Fillers are pronouns, proper names, or (more usually) common nouns that evoke entity or event frames. Entity reference, named entities, and anaphora are all outside the scope of the FrameNet project, but when FEs are filled by frame-evoking words, an interpretation engine should iteratively analyze these words in the same way as any other frame-evoking element.

- Extra-thematic FEs have a considerably different interpretation from all other FE types. Normal FEs (barring Coresets (see Sect. 3.2.2.1) or Excludes relations (see Sect. 3.2.2.3)) must always be logically present for the frame to make sense. Extra-thematic FEs, however, independently evoke a different frame from from the one they are listed in. The Extra-thematic FE itself fills one of the FEs of this frame, and the other FEs are filled by various frame elements of the original target word according to heuristics which must be separately specified for each Extra-thematic FE.

  For example, in 44, evoking the Ride_vehicles frame, the COTHEME FE evokes an additional instance of Motion whose THEME FE is filled by the COTHEME, and whose PATH and other FEs are co-identified with the instance of Ride_vehicles. Thus both "I" and "her" are described as moving to school in this example.

  (44) I **rode** to school [with her COTHEME] all the time.

- Except for the rare LUs with the Semantic Type Biframal_LU (see 6.2.3.3) or LUs in frames which are marked Non-perspectivalized (see 6.2.2.2), all LUs of a frame have a semantics which is best described as a subtype of the semantics of the frame. With these caveats, one may thus consider LU membership in a frame to be an identical relationship to Inheritance from a frame.

6.3.4.2 Inheritance

Not surprisingly, Inheritance, as our most formally defined relation, is an exact match for an ontologically defined relation, namely subtyping or "is-a", as mentioned above. Because these relations are so well-defined formally, there is very little else that needs to be said about Inheritance here.

6.3.5 Propagation of Information

6.3.5.1 Propagation of types on FEs

As discussed above, the Inheritance relation implies the correspondence of the Parent frame and its FEs to the Child frame and its FEs. This straightforwardly allows us to propagate the ontological semantic types (see Sect 6.2.1) on the Parent frame and its FEs downward to the Child frame and its FEs. However, what may not yet be apparent is the fact that the FE bindings in all frame-to-frame relations are equivalent to the Inheritance FE bindings, *regardless of the type of relation.* This means that normal semantic types on FEs

[24] The full interpretation of which frames in a group of subframes are implied to have happened requires the use of X-schema logic or an equivalent system for querying reachability and preceding states in relationally-defined state/event systems. See Narayanan (1999).

can be propagated down the hierarchy along every type of relation. Types are propagated explicitly in the current OWL representation of the FrameNet data.

Because we cannot anticipate all of the Semantic Types that will be useful for tagging FEs, it will certainly also be desirable to categorize the fillers of our FEs using WordNet (or a similar resource). Just as with the pre-marked semantic types applied by the FrameNet team, whatever information is gathered on fillers of an FE in one frame can be propagated to all FEs which are connected to it by any frame-to-frame relation.

6.3.5.2 Propagation of usage information for frame parsing

As shown by Mohit and Narayanan (2003), the frame hierarchy should help in the training of semantic parsers to label sentences with FrameNet frames and FEs. The basic principle is that each FE in an annotated sentence should give some evidence for all the connected FEs, both in terms of the types of fillers expected, and in terms of the expected syntactic realizations. Which of these two types of information for a connected frame can be gleaned from a given annotation set actually depends on which type of frame-to-frame relation we are faced with.

Information on fillers can in principle be gleaned from every type of frame relation (cf. section 6.3.5.1 above), and should be useful for identifying FEs of a target regardless of its part of speech. This means that the annotation of the Theme FE from Taking.seizure.n should help the annotation of the Goods FE with Theft.steal.v.

Since Core FEs may have an idiosyncratic syntactic realization (see the Giving frame), the only way to have absolutely solid information about the realization of the FEs of an LU is to actually consult the patterns seen in the annotation data. However, in practice there is usually a fair amount of overlap in the way that semantically related predicates (especially those of the same syntactic category) realize their arguments. All of this indicates that a statistical parser may reasonably rely on the realizations of LUs semantically similar to the LU of interest as one input to its statistical decision.

6.4 Metaphor in FrameNet

The FrameNet treatment of metaphor is based on the well-known fact that metaphors differ in their novelty/pre-packaging (see Lakoff and Johnson (1980)). Although this pre-packaging is, of course, a matter of degree, FrameNet makes only a binary distinction along this continuum between "productive" and "lexicalized" metaphor, indicated by whether annotation is done with respect to the source domain of a metaphor (the literal frame) or with respect to the target domain (the frame that more directly encodes what the speaker was trying to say) respectively. We indicate productive metaphor by marking with the sentence-level tag "Metaphor" and annotating only in the source domain if:

- All synonyms of the the current target and related terms have a corresponding alternation between literal and metaphorical uses,[25]
- All FEs of the target domain are mapped to FEs of the source domain,[26]
- The sentence can only be understood by relatively consciously evoking the source domain. Although this criterion is especially vague, it has never been used, in practice, as a deciding factor since it correlates so well with the other criteria.

Thus the following examples are labeled as "productive" metaphor:

[25] Often there are semantic differences in instances of the target or related terms that explain their lack of participation in the metaphor, if we take the care to define semantics of the metaphor and the LUs sufficiently specifically. Thus "She **slid** into a coma" is a productive use of the Change of State is Motion metaphor, despite the fact that many of the other targets of the Motion frame (e.g. **roll**, **travel**, **weave**, etc.) cannot be used to describe becoming comatose, and some members (e.g. **circle**) cannot easily be used to describe any change of state. In any frame, there will be LUs with a diversity of semantics, and some specific details of a LU may well block participation in the metaphor, such as in the case of *circle*, which normally evokes a kind of cyclical motion that deemphasizes the changing position of the Theme. Because the change-of-location aspect of motion is deemphasized, *circle* is a poor fit for a metaphor that is entirely concerned with change.

[26] FEs introduced by the metaphor construction are obviously not included in the mapping back to the source domain. For example, in "Why have we chucked all our visions into the **political trashcan** and replaced idealists with policy wonks?", *trashcan*, a member (on the literal side) of the Containers frame, is a valid example of a specific, productive metaphor States of Irrelevancy are Waste Containers (an instance of the States are Locations metaphor), since *trashcan*, *garbage*, *dustbin*, etc. can all be used to convey an Undesirable_situation, even though the modifier *political* does not correspond to an FE of the Containers frame, since *political* represents the Target_domain FE of the metaphor construction.

(45) [Cause_motion] Once our parents are dead, we are **catapulted** into becoming the older generation ourselves.

(46) [Absorb_heat] There was a darkness in his eyes, anger **simmering** just beneath the surface.

These criteria are admittedly themselves vague, but hopefully listing them here helps to clarify what factors are important to consider. So, for example, in sentence 45, it is clear that:

1. *catapult* could be replaced by virtually any member of the Cause_motion frame that has sufficient "force" to convey the same (metaphorically) jarring transition,
2. the FEs Cause, Theme, and Goal are all mapped using the States are Locations metaphor onto Cause, Patient, and State, and
3. the expression is, impressionistically speaking, very vivid.

We see that the criteria are also met in sentence 46, although in this case, against our general policy, we (experimentally!) made the Emotion_heat frame to represent the particular blended semantics of the metaphor Emotion is Heat.

If any of these criteria are not met (i.e. synonyms and semantically related words do not participate in the same metaphor, some FEs are independent of the source domain, or there is insufficiently vivid evocation of the source frame) we consider the metaphorical use to be a separate sense from the literal. So, for example, in sentence 47, there is a frame element that, in the target domain (involving banking) we could call a Bank FE. There is no correspondent to this conceptualization in any putative source domain, as can be seen in sentences 48 and 49. This is because, in a lexically specific way, the source domain's Goal FE is blended with a concept of a person that keeps and preserves, as seen in examples like sentence 50. In 51, we can also see the FE *Variable* which has no correspondent in literal usages like 52.

(47) [Bank_depositing] Bonlat claimed to have £2.7BN **deposited** *with the Bank of America*.

(48) [Alluvial_depositing] *The river **deposited** sand *with the delta*.

(49) [Placing] *She **deposited** her backpack *with the table*.

(50) Brenda has temporarily left her child *with this woman*.

(51) [Position_on_a_scale] It will be so **low** *in price* that no man making a good salary will be unable to own one.

(52) The great dirigible was so **low** that they could see every crease and contour from nose to fins.

In such cases, an LU should be created in a frame corresponding to the target domain of the metaphor, and, in principle, the metaphorical relation between the domains would be modeled as a Frame-to-Frame relation. In practice, however, FrameNet has not yet added any such links.

In neither type of annotation has FrameNet labeled the source and target domains simultaneously, since we deem this to be worthy of a whole research project by itself. For examples of what such a treatment would look like, see Burchardt et al. (2009, 216-9).

Chapter 7

Consistency Management in FrameNet

The quality of a lexical resource such as FrameNet is crucial for its usefulness to both humans and computers. Data quality is evaluated in terms of its consistency and completeness. Consistency means that the body of data should obey any restrictions placed on it and that it should not be self-contradictory. Completeness means that the data set should exemplify frames completely so as to support machine learning and human comprehension.

An example of a consistency requirement is that for a pair of frames related by the Inheritance relation, each core FE of the parent frame is normally[1] mapped to a FE of the child frame. An example of a completeness requirement is that each core FE needs to be exemplified in the frame description. As these examples illustrate, the quality requirements on the data flow from the principles of frame semantics and the desire to support certain kinds of applications. In particular, many requirements are motivated by the theoretical understanding and the envisioned practical applications of Frame-to-frame relations discussed in Chapter 6. Overall, more than 100 quality requirements have been defined. Achieving these requirements turns out to be quite a challenge because the FrameNet data and documentation are maintained continuously and simultaneously by the FrameNet team. In view of the large size of the FN database, strictly manual quality control measures are too costly. Therefore, semi-automatic and, to some degree, formal consistency management approaches are used which reduce the necessary effort considerably. FrameNet's quality management measures are more fully discussed in (Scheffczyk and Ellsworth (2006); Scheffczyk et al. (2006)).

7.1 The Technical Architecture of FrameNet

The center of Fig. 7.1 illustrates the technical basis of FrameNet, which conceptually consists of three databases: The *Lexical Database* contains the relationships of word forms, lexemes, lemmas,[2] and their parts of speech. The *Frame Database* defines and interconnects frames and their FEs. The *Annotation Database* contains annotations and sentences, which comprise the majority of the FrameNet data. FrameNet data are accompanied by documentation – this manual – which makes explicit reference to the databases.

All of these data are connected via the LU table, which associates lemmas with frames and is referred to by the annotation sets. There are many reasons to keep these databases distinct for our purposes:

- Whereas the data in the Frame Database are readily formalizable, most data in the other two databases are less so, since they directly represent the irregularities of natural language word-forms and syntax.
- The amount of data in the Annotation Database is far greater than in the other two databases.
- Whereas the data in the Frame Database are to a fair degree language independent, data in the other two databases are language dependent.[3]

[1] Our handling of requirements includes a formalized concept of exceptions.

[2] A lemma may consist of multiple lexemes.

[3] Nevertheless, the kinds of data in the Lexical Database and the Annotation Database remain conceptually similar across languages, e.g., the concept of multiple meanings corresponding to a single form (i.e. polysemy/homonymy).

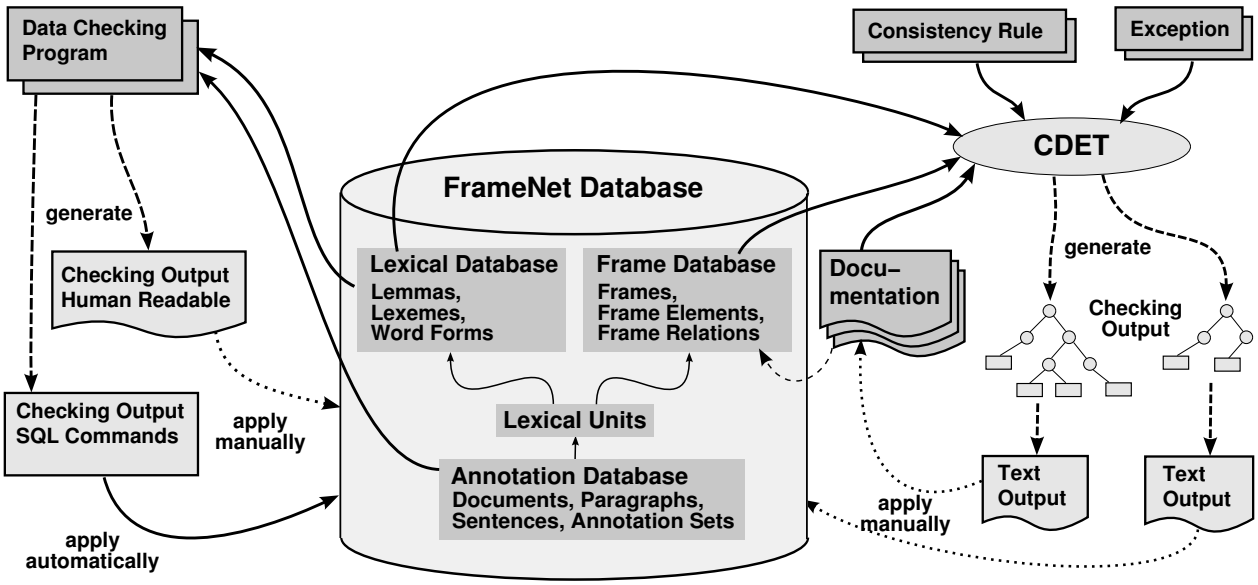

Figure 7.1: FrameNet architecture and quality management approaches

Each of the three databases consists of several tables that are connected to each other. For example, there are separate tables for frames and FEs, where the FE table is linked to the frame table.

7.2 Techniques for Managing Quality

Quality management is is carried out using several different techniques:

1. Some errors are prevented through database structure and native database constraints.
2. Some errors are prevented through restricted database access via a high-level interface.
3. Some errors are tolerated and documented with the help of software tools external to the database and the normal user interface.

The third measure is particularly important because, as experience has shown, violations to quality requirements are inherent to the linguistic enterprise.

Although the high-level interface takes care of many consistency problems, this interface cannot take care of all problems by forbidding inconsistencies: changing or deleting data often violates quality requirements. For example, the descriptions of frames as they appear in the frame report are stored as (XML) text fields in the Frame Database. Within these descriptions, FEs may be referenced. For the description of a specific frame, the interface allows the user to mark up only FEs that are really defined and also belong to this frame. If, however, a referenced FE is deleted or its name changes, these references become invalid.

A violation of a quality requirement might not be an error but an exception to the requirement. In linguistics it is not always possible to fully specify why some data are an exception to a quality requirement. Therefore, FrameNet has implemented approaches to deal with *exceptions*. Common exceptions are test cases (e.g., frames having a name starting with "Test"), which are excepted from consistency checking completely.[4] Providing the ability to document and tolerate errors, at least for the short and possibly for the long term, also lets the FrameNet development team decide whether and how violations are to be resolved.

7.3 Tolerant Quality Management

Fig. 7.1 illustrates our two quality management approaches, which are motivated by the characteristics of the databases.

[4]These test cases exist in the FrameNet databases only and are not part of FrameNet data releases.

A number of data checking programs check annotations in the *Annotation Database* for correctness, completeness, and style. Each program generates a specific error report showing violations of a quality requirement. Thus, the program provides an *algorithmic* (imperative) definition of quality. Depending on whether or not it is clear in advance how a certain type of error can be remedied, the checking tools generate either machine-readable output in the form of native database commands that can be applied to the Annotation Database for automatic repair, or human-readable output for manual inspection.

The chief advantages of imperative quality assurance are fast performance, a very specific output, and the possibility of automatically performing repair actions. With regard to speed of execution in particular, checking programs are the only practical way to check annotations for quality.

For checking the quality of the *Lexical Database*, the *Frame Database*, and *documentation*, FrameNet employs the CDE toolkit (Scheffczyk et al. (2004, 2003)).[5] Here, *declarative* consistency rules define quality in formal logic. For each consistency rule, CDET generates a formal description of violations and possible repairs, which can then be transformed to other output formats.

Defining quality declaratively has a number of advantages: A general-purpose specification language improves the understanding and formalization of quality requirements. It also allows for reasoning about consistency rules. CDET's fairly simple consistency rule language supports incremental consistency checking – a key to tight process integration.

Since the Lexical Database and the Frame Database are sufficiently "relational" and hold a limited amount of data, declarative consistency management can be used. Because of the advantages it offers, we would prefer to use this approach for the Annotation Database also. However, performance issues and, more importantly, the complexity of the annotated natural-language sentences make declarative consistency checking impractical if not altogether impossible.

7.4 Achievements

The application of formal and rigorous quality management has produced both a significant increase in the quality of FrameNet's data and a decrease in the effort necessary for maintenance and repair. The most important quality requirements are now satisfied by the FrameNet data. Violated quality requirements are extensively documented by the imperative scripts and by CDET. The General Release Notes, R1.3, on the FrameNet website contain detailed information about quality requirements and whether they are satisfied or not. In cases of violation, we document which linguistic entities cause the violations.

Although many quality requirements are still violated in the current FrameNet data release, it represents a clear improvement over previous data releases. (1) The most important quality requirements are satisfied. (2) FrameNet and its users have precise knowledge of violations of less important quality requirements. Thus, the quality of FrameNet's data can now be evaluated much more easily by its developers and users.

[5] Consistent Document Engineering Toolkit (CDET), see `http://www.icsi.berkeley.edu/\~{}jan/projects/CDET/`.

Appendix A

Major Extra-thematic Frame Elements

Extra-thematic FEs are ones that combine with many other frames, either taking them into their scope and embedding them in a larger context, or elaborating the descriptions of participants or the setting. Although we have called these entities frame elements, there is a sense in which this is a shorthand. Frame-elements proper identify exactly the parts of a sentence that pick out sub-parts of the scene introduced by a target. In contrast, these elements explicitly introduce an independent scene. There is a very real sense in which the extra-thematic FEs are introduced by separate, constructional targets, which evoke separate frames.

Since, however, the constructions that introduce extra-thematic FEs are so tightly bound together with the structures introduced by regular targets, we believe that it is appropriate to annotate them as if from the point-of-view of the target. In some cases an extra-thematic FE is introduced in a core grammatical position like the object (see, for example, RECIPIENT). Also, an FE that is extra-thematic in one frame may be peripheral or core in another frame. For instance, the frame element ROLE is core in the Replacing frame, but extra-thematic in the Intentionally_create frame. Moreover, several FEs have extra-thematic as well as core and peripheral uses that will be differentiated in future releases. The relevant sub-sections point these out.

To recognize their somewhat independent status, the section that follows lists (1) the meaning and usage of the extra-thematic elements, followed by (2) the structures (syntactic or lexical) that serve to evoke them, and in some cases (3) corrective updates comparing policy at the time of the data release to the current policy (as described in this section).

The list of extra-thematic FEs given here is not exhaustive. Although there are many others, the ones listed here are the most widely occurring. We have in a few cases listed peripheral FEs for comparison. For instance, in section A.1.1, we discuss the peripheral FE DURATION in comparison to the two extra-thematic FEs PERIOD_OF_ITERATIONS and DURATION_OF_FINAL_STATE.

Since some of the entries below describe changes in policy that have not been implemented in time for Release 1.3, we have adopted the convention of marking FEs that are not yet in use with *, and marking FEs that are to be phased out with #. Many of the FEs have not been added to all of the appropriate frames as of yet.

A.1 FEs related to temporal structure

A.1.1 Duration

FrameNet distinguishes several kinds of time span-related frame elements. The basic notion of duration is covered by the eponymous peripheral frame element DURATION.

In Release 1.3, the frame element label DURATION is in several frames still applied to two other time-span related frame elements that are both conceptually distinct from DURATION and extra-thematic, rather than peripheral. We plan to convert a relatively small number of existing DURATION labels to the more specific labels PERIOD_OF_ITERATIONS and DURATION_OF_FINAL_STATE in the release following Release 1.3. Note that the latter two label types iterations already exist in some frames but not in others where they are needed.

A.1.1.1 Duration as a peripheral frame element

This frame element has peripheral status in state or activity frames.

- Meaning: DURATION denotes the length of time from the beginning of a continuous situation (the one denoted by the target) to its end. In many cases, the continuous situation is a dynamic action which is ongoing, while in others it is simply an undifferentiated state.
- Form: PP-through(out), PP-over, PP/Sub-since, enumerated pre-nominal calendric units, or adjectives describing length.

  (1) Cells were **treated** with chemicals [for 30 min] at 37 C or as stated and then incubated for the indicated times before isolation of total RNA .

  (2) I have **known** it [for years].

  (3) They had a [short] **conference** off by the stable.

A.1.1.2 Period of iterations

This FE is minimally distinct from DURATION and DURATION_OF FINAL_STATE in that the clause described is conveyed to be iterative. PERIOD_OF_ITERATIONS can co-occur with either the peripheral FE DURATION or the extra-thematic FE DURATION_OF_FINAL_STATE (although these cannot occur together).

- Meaning: The length of time from when the event denoted by the target began to be repeated to when it stopped.
- Form: PP-for, (often) PP-over, (often) PP-through(out), PP/Sub-since, ADV-"ever since", length-adjectives, pre-nominal calendric units

This FE modifies a clause with an iterative interpretation, which may be signified by the simple present tense on the main verb.

(4) The two writers **talked** about a joint project [for 10 years].

(5) I have been **communicating** with the Minister [since 1988] on that problem.

(6) [For many years], he **walked** to the forum alone.

A.1.1.3 Duration of final state

- Meaning: DURATION_OF_FINAL_STATE denotes the length of time from the beginning of a state resulting from the activity denoted by the target until the state no longer holds.
- Form: PP-for, (not PP-over except with specific heads like "weekend"), (rarely) PP-through(out), (never *since*), enumerated calendric units, and length-adjectives.

  (7) The previous day President Ranasinghe Premadasa had refused to extend the [seven-day] **cessation** of hostilities which the government announced on Jan. 3.

  (8) If you **place** it there [for extended periods], your monitor may suffer permanent damage.

A.1.2 Frequency

- Meaning: This frame element is defined as the number of times an event occurs per some unit of time. A FREQUENCY expression answers the question *how often*. It is to be distinguished from ITERATION, which pertains simply to the number of times an event occurs—ITERATION expressions answer the question *how many times*. In the future, we will further distinguish PARTICULAR_ITERATION, which is intended to pick out a particular iteration from a series. The annotation in Release 1.3 is not yet consistent in making these distinctions; the frame element PARTICULAR ITERATION so far exists in only a few frames.
- Form: A calendric-unit plus *-ly*, *every* + calendric-unit, *time(s) (per)* + calendric-unit

  (9) Lilly Foley never missed her [weekly] **hair-do** and manicure .

  (10) The scribes **copied** the scripture [every 80 years].

  (11) Those affected may be **vomiting** [many times a day] and may also be abusing laxatives .

- Meaning: The frame element ITERATION is used for expressions that indicate the number of times an event or state (of the kind denoted by the target in its clause) has taken place or held.
- Form: *once*, *twice*, Cardinal expression + /it times, /it lots/a lot

(12) The ferry that Kenneth was on was **hijacked** [twice].

(13) Clashes **broke out** [several times] because Kenneth had been arrested.

(14) An Agoura Jills cigar store was **robbed** [twice] last week.

(15) Earlier, Private Lee Clegg (22) told the court he had **fired** [three times] in the vicinity of the car driver's head.

A.1.4 *Particular iteration

- Meaning: Expressions marked with this extra-thematic FE modify a non-iterative use of the target, and indicate that it is conceived as embedded within an iterated series of similar events or states. In addition, most expressions of PARTICULAR_ITERATION indicate which instance of the series is being referred to.
- Form: *again*, *(for) the/a _th time*

(16) Four of the five protagonists of this story – Greco-Macedonians , Romans , Jews and Celts – **came together** [for the first time] in the Hellenistic period.

(17) She really ought to **rehearse** that presentation [again], she thought.

A.2 FEs related to places

A.2.1 Location of protagonist

- Meaning: This frame element is used to indicate the location of a participant in cases where the multiple participants of an event are not co-located in the typical fashion. In Release 1.3, this FE has various frame-specific names including LOCATION OF PERCEIVER, or LOCATION OF PARTICIPANT; it was even (incorrectly) labeled PLACE in some frames. This frame element is attested in our data mostly with perception and communication events, as shown in (18)-(21). However, the label is applicable to a much wider range of event types (cf. (22)-(23)).
- Form: PP-from, including cases of "preposition-stacking" (24).

(18) No deaths or injuries, but a few hands **waving** feebly [from under the pile of bodies] indicated that rescue was required. (Currently, examples of this sort are wrongly annotated as Source in the Body movement frame!)

(19) The cottage still **looks** very much the same [from the outside].

(20) The witnesses **saw** the robbery [from the their car].

(21) Does anyone **e-mail** the office [from a phone booth]?

(22) Smithers **finished** his work [from home].

(23) He was suspected of **running** a business [from behind bars].

(24) I couldn't believe that he could **hit** me [from between the goal posts] !

A.2.2 Subregion

- Meaning: SUBREGION denotes a sub-part of an event participant in which the state or process denoted by the target more specifically holds or unfolds. Note that this FE is only usable in cases where the target involves some property which is relevant for subparts of a participant. The target may predicate a change of such a property or itself denote a state of the property. In the case of physical objects, it is most often used with directly confirmable characteristics like temperature, color, consistency, visibility, etc. It can also be used of abstract entities (30).

- Form: Stative locative expressions

  (25) The mousse was **chilled** [around the edges].

  (26) The cloth was **dried** [around the edges].

  (27) All the walls were **black** [on the inside] with fluorescent pink skylines of places like New York painted on.

  (28) The gazebo was **screened** by a hedge [on one side].

  (29) He didn't **paint** it [on the bottom].

  (30) This article is **terse** [in the beginning].

  (31) The batlith is a bit **rough** [between the handles].

A.2.3 Point of contact

- Meaning: This FE is used when there is some part of a participant that makes contact with a horizontal surface. In some cases, it denotes a body part that serves as the support for a body on a surface, while in other cases it refers to a part of a theme that makes contact as a result of motion.

- Form: PP-on(to)

  (32) The horse **pitched** the rider [on his head].

  (33) Quilp had shut himself in, **stood** [on his head] before the door, then walked on his hands to the back and stood on his head there.

- Update: This frame element should also be prominent in the Motion_directional frame. However, there are no annotated instances in that frame in Release 1.3, as the frame element was defined for that frame only after the vast majority of the annotation was completed.

A.3 FEs related to additional participants

A.3.1 Recipient

- Meaning: The label RECIPIENT is used for actual or intended recipients of some theme. Generally the agent of the clause headed by the target instantiates the donor, and the theme is also a core argument of the target. This FE should be compared to BENEFICIARY, which describes those who benefit from an action but are not participants in a transfer scenario.

- Form: Object, PP-for

  (34) Bill **baked** [Kim] a cake for Arbor Day.

  (35) Can you **knit** [me] a sweater?

  (36) The customer had **bought** a book [for his sister].

A.3.2 Beneficiary

- Meaning: This extra-thematic FE applies to participants that derive a benefit from the occurrence of the event specified by the target predicate. Further, the target predicate should involve some sort of agent that intends that the benefit go to the BENEFICIARY.

- Form: PP-for, (dialectally) Object

  (37) Someone even **mowed** the lawn [for Camile and her family].

  (38) My cart arrived yesterday and my husband **assembled** it [for me] today.

A.3.3 *Maleficiary

- Meaning: This frame element is defined as the counterpart of Beneficiary, and is used for participants that suffer a harm or loss from the occurrence of the event specified by the target predicate.
- Form: PP-on

  (39) Every pet I've had has **croaked** [on me].

  (40) The volume button **broke** [on me] the first time, so I returned it and got a new one.

Update: This Frame Element is currently not defined in any frame but will be used in future releases.

A.3.4 Co_participant

- Meaning: Co_participant is an entity that participates in a coordinated way in the same event as the primary protagonist, regardless of whether the protagonist, and hence the Co_participant, is more agent- or more undergoer-like.
- Form: PP-with, PP-along with, PP-together with

  (41) Self_motion: When we were hunters, two and a half million years ago **roaming** round [with the monkeys], only 15-20 per cent of our diet was fat.

  (42) Departing: He had **emigrated** from Hampshire, [with 26-year-old girlfriend Tracey Farmer] to escape the recession and start a new life.

  (43) Placing: Pat **put** the butter in the fridge [along with the margarine].

  (44) He was **arrested** [with his brother] on 8th November 2002.

  (45) [Along with two assisting officers], Smithers **arrested** all twenty of the much-sought traffickers.

- Update: In Release 1.3, the label Co_participant has not yet been extended to all the instances of this semantic notion. Many of them are still incorrectly covered by the extra-thematic FE Cotheme.

  Although the uses of Co_participant (formerly Cotheme) that exist in the database at the time of this writing are concentrated in motion-related frames, the frame element Co_participant is intended to be used more generally for all kinds of events. In the Arresting frame, for instance, the extra-thematic frame element Co_participant could be used to label phrases denoting people that are arrested (44) along with the Suspect, or others that assist the Authorities (45).

A.3.5 Cotheme

- Meaning: This frame element is used to denote phrases that specifically entail an entity which is conceived as moving relative to another entity. It is mostly used in several motion frames, often describing the motion of something that responds to how a Self-mover) moves.

  Occasionally, Cotheme can be used in non-motion frames, in cases when something is conceptualized as following a mover. Although the conditions for such conceptualization may be broader in other languages, in English this is usually limited to cases where a speech signal is conceived as following a moving Addressee.

- Form: PP-after, PP-ahead of

  (46) Self_motion: "Come," he said, and the rabbit **hopped** [after him].

  (47) Self_motion: She became highly indignant if we **walked** [ahead of her].

  (48) Communication_manner: She **shouted** angrily [after him], but he paid no heed.

- Update: The many uses of this frame element in this release which are instantiated by a PP-*with* belong properly to the Co-participant FE.

A.4 FEs describing participants

A.4.1 Role

- Meaning: This FE identifies the role filled by a particular participant of the clause headed by the target. In so doing, it relates the state of affairs evoked by the target to another background state of affairs by indicating that one of the role fillers of the current frame also fills a second kind of role in the background, either generically or in some specific instance. As is the case with many Extra-thematic and Peripheral FEs, this FE may require second-layer annotation as Explanation (50).

- Form: PP-as

  (49) All of which was done without **informing** me [as your lawyer] which has now resulted in contempt of court proceedings being instituted against you.

  (50) [As your friend], I **demand** that you go out on stage and wow them like I know you can.

A.4.2 Depictive

- Meaning: This FE describes a participant of the state of affairs introduced by the target as being in some state during the action. The depicted state does not necessarily or usually facilitate or cause the state of affairs reported by the target. In some cases, the Depictive describes the participant by characterizing a subpart (51-52).

  When the depictive phrase both depicts a participant and gives rise to an inference that the depicted state enabled or facilitated the state of affairs reported by the target, we annotate Explanation on the second FE layer (58-60).

- Form: quite variable, including AJP, PP (esp. *with*), SAbs, VPing; the position of the phrase is clause-initial, immediately post-subject, or (among many other adjuncts) after the object and particles (if present)

  (51) Smithers **stormed** out of the office, [his temples throbbing with rage].

  (52) Jess grabbed the bag and **walked** out [with her head held defiantly high].

  (53) [Weighing not much more than a portable at 6.5 pounds], the LT260 also **features** 2100 ANSI lumens and sharp XGA resolution.

  (54) Sue **came** in [holding a black candle in her right hand].

  (55) "One of the most common things volunteers will say when they come back is that they went over to help and teach, but **came** back [being the ones that learned the most]-they were the ones being taught."

  (56) Coming into it, I think I was in a tough situation because I probably had more to lose than I had to gain. I **came** in [being the top kicker], I came in making 90 percent of my field goals in the last two or three years, so I came in with a lot to lose.

  (57) Brynda **fell** back, [her face pale with exhaustion and pain].

  (58) [Wearing my feather cap] I was **spotted** immediately by Bill.

  (59) [Growing nervous due to the late hour, and not having heard any sounds emanating from the room], the secret service agents **rap** on the door.

  (60) [Being a wittarian], John **had to** bite his words before speaking.

A.5 FEs describing events

A.5.1 *Event_description

- Meaning: In general, this FE is used for phrases that describe the clause of the target as a whole. There are two somewhat distinct subgroups of sentence modifiers that cover this function.

First, it applies to appraisals of the event, typically a judgment of its expectedness (61-62). The notion of appraisal requires a factive portrayal of the event.[1] Adverbs such as *likely, presumably, possibly* etc. are not covered by it since they specify a non-factive understanding.

The second subgroup consists of any descriptions of the state of affairs denoted by the target as role fillers in other frames. The state of affairs reported by the target is understood to play a role in the state of affairs of the modifier. What that role is depends only on the predicate heading the modifier phrase. Typically the modifier takes the form of a finite sentential relative clause, as in (63) and (64), or of a VPing, as in (65).

EVENT_DESCRIPTION should be differentiated from RE_ENCODING, characterized below.

- Form: *ly*-adverbs; VP-ing, non-restrictive Relative clauses

(61) [Amazingly], the WiX installer is **turning out** to be almost entirely pleasant.

(62) [Oddly], **there was** a very small earthquake in Miami today.

(63) I haven't **found** much of use in MSDN, [which surprised me].

(64) In the first ever live Soccer AM ping-pong championship, Tim **beat** Gazza 11-6, [which was revenge for his previous 10-0 defeat]!

(65) Awaiting execution, Peace finally **confessed** to the murder of PC Cock, [providing the detailed evidence to persuade the authorities of Habron's innocence].

A.5.2 Re_encoding

- Meaning: This FE presents the current frame as an integral part of a larger conceptualization expressed by another frame. The frame expressed by the target typically corresponds to a MEANS action within the frame expressed by the RE_ENCODING, but not necessarily, as is shown by (67, 68, 69).

  A more specific sub-type of the RE_ENCODING idea is illustrated in 73, which specifically presents the current event as an intentional act undertaken with the purpose of bringing about the state of affairs expressed by the other frame.

- Form: NP, PP-in, PP-as, PP-for

(66) A large piece of the glacier **fell** suddenly into the sea, [a stunning display of Nature's might].

(67) [In a shocking turn of events], the free iPods pyramid scheme **turns out** to be a pyramid scheme.

(68) [In an uncanny quirk of fate], he and his current dog were **visiting** the animal hospital the same day I found it necessary to say a final "Farewell" to Dudley.

(69) A large piece of the glacier **fell** suddenly into the sea, [in an amazing display of Nature's might].

(70) [In reprisal], the BDR and the local Muslim settlers **raided** Sukhnachari.

(71) Justin Hawkins, frontman for The Darkness, **pierces** himself [as punishment for his own transgressions].

(72) [As a result of the war], a growing percentage of Muslims **see** serious threats to Islam.

(73) If you **send** people home [for punishment], let them have their pay at least.

A.5.2.1 Differentiating Re_encoding from Event_description

RE_ENCODING phrases differ from those of EVENT_DESCRIPTION in that with the former the state of affairs of the target **constitutes** the state of affairs of the modifier. Constituents labeled as RE_ENCODINGs allow paraphrases of the form "Event of modifier *consists of* Event of target". For instance, (70) can be paraphrased as in (74):

(74) The reprisal *consisted of* the BDR and the local Muslim settlers raiding Sukhnachari.

[1] When appraisal adverbs are embedded in non-factive contexts such as conditionals, the overall interpretation of the sentence may be non-factive: *If, [surprisingly], it is NOT me you are trying to woo, then just say this same thing to whomever you ARE trying to woo.*

Some EVENT_DESCRIPTION phrases allow similar paraphrases but in principle EVENT_DESCRIPTION phrases need not allow them and very often they do not. While (75) may be an adequate paraphrase for (64), above, (76) is not a good candidate paraphrase for (63).

(75) His revenge for his previous 10-0 defeat consisted of beating Gazza 11-6 in the first ever live Soccer AM ping-pong championship.

(76) ?*My surprise consisted of not finding much of use in MSDN.

We maintain that the acceptable paraphrase in (75) is just an accident of the fact that the predicate of the sentence modifier in (64) is *revenge*, a noun which can take a MEANS phrase as subject in copular clauses. (76) is not a possible paraphrase for (63) because the noun *surprise* (in contexts of unintentional surprising) only takes a CAUSE rather than a MEANS frame element as subject.

The distinction between EVENT_DESCRIPTION and RE_ENCODING is particularly clear in the case of relative clauses. This formal type occurs only in EVENT_DESCRIPTION phrases and it allows for the two events to be completely distinct, which is not possible with RE_ENCODING FEs. Thus, (77) involves an EVENT_DESCRIPTION and no reformulation with a RE_ENCODING construction is felicitous (cf. 78).

(77) Thereafter she **became** pregnant, [which was discovered by her parents on July 24, 1941 EVENT_DESCRIPTION].

(78) *[In the discovery by her parents on July 24 1941] she thereafter **became** pregnant.

A.6 FEs related to co-occurring events and circumstances

A.6.1 Circumstances

- Meaning: CIRCUMSTANCES describe the state of the world (at a particular time and place) which is specifically independent of the event itself and any of its participants. This breaks down into two major uses: phrases expressing prevailing physical conditions, as in (79)-(83); phrases expressing concurrent states of affairs which are neither RE_ENCODINGs (that is, larger events of which the event expressed by the target is part), nor properties predicated of a participant of the frame evoked by the target. Examples of this are found in (84)-(87).

- Form: PPs, esp. *with, at, under*

(79) Not just red light, but dim red light. < 3 lux, and you have to be sure not to directly shine it into an animal's eye. I've **worked** [under those conditions] for years.

(80) Place the apples into a greased and floured pan and **bake** [at high heat].

(81) [Under those weather conditions], the trapped food will **form** brilliant purple and red chemicals in the leaves.

(82) [In this weather] the climb was very **difficult**.

(83) His car caught **fire** [at that temperature].

(84) She **entered** in grand style [with everybody looking at her].

(85) They **continued** [with Chris in charge of the hockey club]

- Update: Examples like (86), in its context, have causal implications. Formerly we used the CIRCUMSTANCES label in particular for constituents that introduce factors which enable or facilitate the occurrence of the state of affairs reported by the target. Now we no longer require the presence of such causal implications to use the label CIRCUMSTANCES. Instead, we use second-layer annotation to indicate the contextually enhanced understandings of such phrases. Thus, while (86) would carry a second-layer EXPLANATION label, we would apply a second-layer *CONCESSIVE label to (87).[2]

  In Release 1.3, we still have a relatively rare FE called LEGAL_BASIS, which we now take to be subsumed under CIRCUMSTANCES. Just like other examples of CIRCUMSTANCES, it may be used with or without causal implications; cf. (88) and (89).

[2]For a discussion of how to test for special causal, concessive, and conditional interpretations of TIME, PLACE, etc., phrases, and of the linguistic contexts that trigger such interpretations, please see section (A.9.4).

(86) This program will not **survive** [with everybody going in different directions].

(87) What was it like **going** into Bastogne, [with everybody going the other way]?

(88) We are all **living** [under Californian law].

(89) His assets were **seized** [under Section 9.3 of the Uniform Code].

A.6.2 Containing_event

- Meaning: This FE denotes an event that occurs or state of affairs that holds at a time that includes the time during which the event or state of affairs reported by the target occurs and of which it is taken to be a part.

  As (91) and (92) show, CONTAINING_EVENTs are frequently understood to have a facilitating or enabling relation to the event expressed by the target. A causal interpretation is, however, not strictly necessary, as shown by (90). When it is present, it is recorded with second layer annotation.

  CONTAINING_EVENT does not simply locate an event relative to another known event. (See below.)

- Form: PP-in, PP-on, PP-at. Most CONTAINING_EVENTs are expressed by *in*-PPs but for certain kinds of CONTAINING_EVENTs other prepositions are used. When the CONTAINING_EVENT is a motion event, *on* is used (93). *On* is also used for certain kinds of iterated sub-events such as *turn, shift, leg* etc. (94). *At* occurs as a marker of Containing_events that involve interpersonal interaction (95).

(90) Every single thing I **said** [in the election campaign] about interest rates was right.

(91) [In the wake of the war], a growing percentage of Muslims **see** serious threats to Islam.

(92) Two people were **killed** [in the accident].

(93) I **passed** through this village [on my journey to Castleton].

(94) I recall Ron coming home with red hots which I detested, so [on my turn] I **picked out** some candy shaped like pork chops, potatoes, and peas.

(95) Printed materials will not be **distributed** [at the meeting].

A.6.2.1 Differentiating Containing_events from Re_encoding

The CONTAINING_EVENTs that formally are *in*-PPs may appear similar to RE_ENCODING FEs. They are, however, a separate notion since they do not allow a paraphrase saying that the event denoted by the target fills a role in the event expressed in the extra-thematic phrase.

(96) His heart failed the next day and somehow, this letter **got** lost [in the turmoil].

(97) *The turmoil was that this letter got lost./The turmoil consisted of this letter getting lost.

A.6.2.2 Differentiating Containing_events from Time

The label CONTAINING_EVENT is different from simple TIME expressions (98- 101) whether they anchor an event relative to another event or to a clock time. Even when TIME FEs make reference to an event, this event is considered purely in terms of when it took place. As such, the expressions involving events could be adequately understood if the events were replaced by the time periods in which the events took place; purely clock-time re-phrasings of CONTAINING_EVENTS are clearly not equivalent, however. Cf. (102) and the non-equivalent (103).

(98) You can do **laundry** here [during your stay].

(99) We **pray** [before the meeting begins].

(100) The ice sculpture **melted** [while it was under hot water].

(101) [At 8 o'clock], the butler will **serve** tea.

(102) Twenty people were **injured** in the avalanche.

(103) ?Twenty people were **injured** between 7:05 and 7:07 AM.

A.6.3 *Coordinated event

- Meaning: The label COORDINATED_EVENT is to be used for phrases denoting an event–it does not allow states–that the event expressed by the target is rhythmically aligned with. The COORDINATED_EVENT is conceived as independent: it would occur regardless of the event expressed by the target, which is not even an incidental or optional sub-part of the COORDINATED_EVENT. The COORDINATED_EVENT also is not a RE_ENCODING of the one denoted by the target. Alternatively, the COORDINATED_EVENT can be metonymically referred to with nominals headed by *rhythm, beat* etc. The frame element COORDINATED_EVENT is particularly frequent with motion events of one kind or another, as in (104)-(105), but it also occurs with other kinds of events (cf. (106)-(107)).

- Form: PP-to

  (104) Carnaval '78 packed in over a thousand bodies, sweating and **gyrating** [to the drumming of Austin's first Carnaval group].

  (105) She **danced** [to the drumming of ancestors and the rattle of elders].

  (106) Hearts **Beat** [to the Rhythm of Biological Clocks]

  (107) Lily's best smile fades as she quietly **eats** her dinner [to the hum of Daddy's voice on the phone].

A.6.4 *Correlated_variable

- Meaning: The label CORRELATED_VARIABLE is to be used for phrases denoting a scalar property that can change over time and whose movements are tracked by another variable property expressed by the target (or the target and its dependents). In (108), for example, the amount of knowledge on cancer prevention increases proportionally with time. The CORRELATED_VARIABLE is typically understood as an independent variable while the variable property expressed in the target clause is taken to be a dependent variable. It is not linguistically necessary (nor scientifically correct) to infer causation from correlation, which is shown by (110). In context, however, CORRELATED_VARIABLE expressions are often meant and understood as expressing causes. When the content and context of the target's clause suggests that this is so, we also record this fact with an EXPLANATION label on the second layer of annotation.

- Form: PP-with

  (108) [With each passing month], we **discover** more and more about how to prevent cancer.

  (109) The most recent study found that Okinawans eating a more traditional diet did not **gain** weight [with age].

  (110) A country's standard of living **increases** [with the educational attainment of women]–or is it the other way around?

A.7 FEs related to the causal chain

A.7.1 Explanation

- Meaning: The EXPLANATION denotes a proposition from which the main clause (headed by the target) logically follows. This often means that the EXPLANATION *causes* the state of affairs expressed by the target, but not in all cases. In particular, there are cases where the connection is a bit weaker than normal for causation, as in (111). In these instances, cause is mediated by some actor responding in a complex way to the situation; such a sentence is not at all paraphrasable in the language of direct causation (112). The FE EXPLANATION is also used on the second layer of annotation to indicate that a particular frame element is construed as causal or enabling, as in (113-116), each of whose first layer FEs are as indicated. (See A.9.4.)

  The label EXPLANATION also covers expressions such as those in (117) and (118). These assert that the explanation of the event is something like "for no reason that could be expected" or "because of a mistake".

(111) The lavender is **in** the fridge [because the packet said to sow them in 30-40 degree temperatures].

(112) The packet's saying ... *caused* the lavender to be in the fridge.

(113) [In the nebula Place], you **move** really fast .

(114) Plants **grow** really well [when you pay attention to them Time].

(115) The crowded office was **empty** [with both of the other occupants gone Circumstances].

(116) [Wearing my feather cap Depictive] I was **spotted** immediately by Bill.

(117) In the event your PPT file gets **erased** [on accident], the computer crashes, or the projector doesn't work, be prepared to go on with your talk without it.

(118) He worked there for several years, **becoming** an author [almost by chance].

- Update: This frame element replaces a number of frame elements (Reason as an extra-thematic FE, Cause as an extra-thematic FE, and some instances of Internal_cause and External_cause) whose old uses are detailed below. The new frame element Explanation is not implemented as of this release.

A.7.2 FEs replaced by Explanation

A.7.2.1 #Reason

In most cases, the frame element Reason has been applied to constituents expressing a state of affairs to which the Agent is responding in performing the action indicated by the target.

(119) Listen, I just **mug** people ['cause I got mouths to feed] – nothing personal.

(120) North was **dismissed** [for gross mismanagement of funds].

In some frames, the frame element label Reason has been used for situations involving more immediate causation. The verb *decay* in example (121) belongs to the Rotting frame, which does not involve an agent-like participant. The adjective *tired* in (122) belongs to the Biological_urge frame, which concerns internal experiences of one's body state that are not under conscious control. In (123), which is from the Sounds frame, the appropriate frame element to use is External_cause, about which see below.

(121) Typically, the front top six teeth will **decay** because of the way the child has sucked on its bottle.

(122) Jen is **tired** [because she gave birth last night].

(123) Brian **let out** a hoarse screech [at each cut].

There is a further use of the FE Reason, with PPs headed by *for* describing an activity which is to be balanced by the action of the main clause, as in (124). Such examples will be annotated with the new FE *Reciprocation.

(124) The DEA **took** his license away [for prescribing large quantities of opiates to his chronic pain patients].

(For other deprecated uses of Reason, see Section A.9.1 below.)

Update: Although frames newly created since Release 1.2 no longer use an extra-thematic FE Reason but instead feature the now preferred Explanation label, there are still many frames in which the Reason frame element has not been converted to Explanation.

A.7.2.2 #Cause

This frame element has been used to introduce a causing event or enabling state that causes a state to hold or an inchoative event to occur. The FE is also used for events or states that prevent other states or events. In some frames, Cause also covers entities that are centrally involved in events or states causing, enabling, or preventing other events or states.

(125) Smithers **became** tired [from all the work he'd been doing].

(126) The mop is **drenched** [from the water you spilled].

A.7.3 *Excess

- Meaning: This FE indicates that the presence of an overwhelming amount of a kind of entity or the high intensity of an event is sufficient to bring about the state of affairs indicated by the target predicate.
- Form: PP-with

(127) I was **doubled** over [with uncontrollable fits of laughter] as I contemplated the idea of Kyle being non-partisan.

(128) Everywhere there lay scattered arms, corpses, and mangled limbs, and the earth **reeked** [with blood].

(129) Blackbeard **shook** [with laughter].

(130) In the afternoons, the Sirius booth **rocked** [with live performances] from musicians of all genres.

- Update: Constituents properly tagged as EXCESS are currently tagged with a number of different FEs. For example: although currently tagged as CAUSE, (130) and (129) from the Move_in_place frame should be assigned to the FE *EXCESS. Similarly, instances of the FEs INTERNAL- and EXTERNAL_CAUSE that consist of PPs headed by *with* should also be re-assigned to the FE *EXCESS.

A.7.4 Causal FEs based on Emotional and Cognitive-states: Internal and External Cause

In theory, frames that had an FE INTERNAL_CAUSE should also have FEs called EXTERNAL_CAUSE. However, the frames with an INTERNAL_CAUSE FE defined are much more numerous than those which have an EXTERNAL_CAUSE FE defined, even though at some point we had the notion that the two FEs should always be paired. INTERNAL_CAUSE frame elements are found in: Body_movement, Communication_noise, Communication_response, Facial_expression, Hostile_encounter, Judgment_communication, Make_noise, Making_faces, Questioning, Renunciation, Reveal_secret, Self_motion, Sounds, and Statement. EXTERNAL_CAUSE FEs occur in these frames: Body_movement, Facial_expression, Making_faces, Self_motion. Some expressions like (131)-(132) look like EXTERNAL_CAUSE or INTERNAL_CAUSE but consist of *from*-PPs, and should be annotated as EXPLANATION:

(131) I **shivered** [from the cold].

(132) He let out a deep, hungry growl as she bucked and **gasped** [from the shocking sensation], her eyes like a frightened doe's.

PPs headed by *from* should be annotated as EXPLANATION, given the formal and functional similarity of the construction below:

(133) She **died** [from malaria].

A.7.5 Internal_cause

- Meaning: This special type of cause-related FE requires that a sentient be affected by some internal state so that they bring about the state-of-affairs conveyed by the clause of the target. INTERNAL_CAUSE frame elements are typically realized as prepositions headed by *in* or *out of*. For INTERNAL_CAUSE expressions headed by *in*, there is an additional implication that the internal state of the sentient is perceivable to others because of the event mentioned in the main clause. *Out of* PPs also specifically refer to emotional states, but they do not require that the main clause somehow signify the emotional state. *Out of*-PPs are thus usable with a wider array of predicates than *in*-PPs are, as can be seen from the contrast between (138) and (137). Despite this contrast, we treat examples like (138) as cases of INTERNAL_CAUSE.
- Form: PP-of, PP-in, PP-out of

(134) She checked a **giggle** [of hysteria].

(135) Kim **frowned** [in concentration].

(136) Kim **threw** her hands up [in despair].

(137) ??She **stole** his car [in anger].

(138) She **stole** his car [out of anger].

(139) I was beginning to think that I was the only person to see a horse **kick** [out of anger].

- Update: In earlier annotation practice some *from*-PPs were treated as INTERNAL_CAUSE. However, from now on, we simply consider all of them instances of the EXPLANATION FE since, while they are compatible with internal experiences of an AGENT, this is not required, as shown by the contrast between (140) and (141).

(140) Stumpy tail wagging , the wire-haired terrier **trotted** [from force of ingrained habit] to the back door where his lead hung on a hook.

(141) In the past, a narrow path along the cliff led to the cave, but it **collapsed** [from erosion and rock falls].

A.7.6 External_cause

- Meaning: The FE EXTERNAL CAUSE, also requires a COGNIZER: it denotes an event or state that causes the actor to act, although not in a conscious response to the EXTERNAL CAUSE.
- Form: PP-at

(142) They **danced** [at the sound of her voice].

(143) Andreyev **smirked** [at the old man's passion].

(144) As with Ben Johnson most South Africans **smirked** openly [at these revelations] .

(145) Glancing up, she caught Ross's nod and warm **smile** of approval [at the efforts she was making to reassure Emma and Sophie about their parents' condition].

- Update: There are some annotations with this label in the database that we have reanalyzed. (146) should primarily be treated as DEPICTIVE. (147) is better analyzed as a COORDINATED_EVENT, similar to *Film composer Maurice Jarre, white-dressed,* ***enters*** *the Berlioz Opera Hall [to great applause].* (148) and similar expressions such as *by chance, coincidence, luck* etc. are now treated as cases of the more general FE EXPLANATION.

(146) [Spotting a thrush on the lawn], the cat **leapt** down and gave chase.

(147) Noreen O'Neil looked good as she **pranced** across that stage [to the catchy Irish tune].

(148) [By a lucky accident] we had already **trekked** (with Mohammed) from Imlil westwards to the Nfis, discovering superb trails through a bold landscape.

A.7.7 *Concessive

- Meaning: This FE signifies that the state of affairs expressed by the main clause (containing the target) occurs or holds, and something other than that state of affairs would be expected given the state of affairs in the concessive clause. In other words, given only the facts of the *CONCESSIVE, one would expect them to cause the world to be the opposite of what is portrayed in the main clause.
- Form: a clause or phrase headed by *although, (even) though, despite, notwithstanding, nevertheless*, etc.

(149) Many teachers **favor** charter schools [although their unions don't].

(150) [Nevertheless], it would have been **interesting** had he won in '92.

(151) [A lack of witty sayings notwithstanding], I **can** Safely say that i am currently appalled by the prospect that one's ears and nose may continue to grow throughout adulthood.

(152) Movie industry revenue is **booming** [despite the current level of piracy].

(153) I **received** a confirmation and order number the same day, [despite the store being closed for religious holidays].

(154) [Although no longer a minister], Smithers still **officiated** at weddings.

A.7.8 *Apparent_conclusion

- Meaning: This FE describes an addition to the main-clause that would seem to hold given the rest of the main clause. Quite often, the *Apparent_conclusion contains another FE, either peripheral or extra-thematic, which is non-factive (i.e., it may, but does not necessarily, hold), as in (155)-(157). For these cases, we annotate the other frame-element on the second layer; thus, e.g. (155) should be annotated with Purpose on the second layer. In some cases, it marks a completely separate conclusion than one would draw from the manner of the event in the main clause, e.g. (158).

  (155) [As if to insult our intelligence further] each joke is **followed** with the words, "get it" just in case we don't.

  (156) We observe ourselves '**doing** the act', [as if on stage].

  (157) We find ourselves **making** the point articulate [as if for another listener than those in front of us].

  (158) You have to **negotiate** [as if relationships matter].

A.7.9 Reciprocation

- Meaning: This FE denotes an event paired with that of the main clause, which explains why the event of the main clause is expected. In particular, *Reciprocation is used to indicate an action that is being rewarded or punished in the main clause, or an action that is paired with the action in the main clause as the other side of an agreement.

- Form: PP-for, PPing-for

  (159) I had to stay up and **read** all night [for my sins of previous omission].

  (160) He **gave** me a new watch [for mowing his lawn].

Update: The FE Reciprocation is used in a handful of frames as of Release 1.3.

A.8 Other

A.8.1 Degree

This frame element selects some gradable attribute and modifies the expected value for it.

(161) That photo of the prime minister is [very] **nice**.

Some uses of the label Degree are ambiguous between just specifying a value for an attribute that holds for a particular participant and specifying how much of a group or a single entity participates in the specified state or relation.

(162) The medication should have [fully] **cured** him.

Example (162) may specify the degree of recovery to full health for the whole person or it may specify how much of the person has been fully cured. Consider the continuations in (162') and (162").

(162') But it didn't. He still has some slight pains.

(162") But he still has pains in his legs.

A clear case in which Degree is used to refer to the involvement of a so-called "incremental theme" in an event is (163).

(163) My uncle likes to [fully] **immerse** cookies into his coffee.

Other uses such as (164) and (165) below exhibit a related but slightly different construal. In those examples, the scale that is involved is the proportion of a group (figuring as a frame element) that participates in a specified state or relation. These uses are comparable to cases of quantifier float such as (166). However, they cannot replace all instances of quantifier float with the same meaning. We will use the label DEGREE for cases like (164) and (165).[3]

(164) He **replaced** the 6761 models [entirely] with about half as many 6800s.

(165) In the 1950s the municipality bought 87% of the shares and it **owns** them [entirely] today.

(166) This is the third pet he has lost in a two year span (we **acquired** them [all] when the kids were young).

Another group of uses consists of cases where the prototypicality of a specific event or relation for its type is evaluated.

(167) There were unsubstantiated allegations of [serious] **misbehaviour** while Scotland were in Berne for the first World Cup qualifying Group I tie against Switzerland.

(168) The townspeople engaged in a [full-scale] **free-for-all**.

We also keep uses like (167) and (168) within the range of what the frame element DEGREE covers.

In the annotation of Release 1.3, there is still some confusion of DEGREE frame elements with FREQUENCY, which is in particular due to expressions such as *a lot*, *a little*, and *little*. Consider the following examples:

(169) I've carved a butter knife with it, spoons, built shelters, feather sticks, prepared dinner, chopped down trees, split wood and **sharpened** it [a lot].

(170) Your car **breaks down** [a lot].

(171) My cat **sleeps** [a lot].

(172) He **talks** [a lot].

(173) So, he says, he **sleeps** [little]. "I follow the words of Napoleon," says Shen. "Anyone who sleeps more than four hours a night is a fool."

Examples (169)-(170) are readily understood as saying that events of a certain type happen frequently.[4] Examples (171)-(173) are ambiguous: they could either be understood as indicating that an event happens frequently or that the specified events go on for a long overall duration. The specified duration may however be achieved over just a few occasions: the cat in (171) may sleep for 12 hours straight every day, but then be awake without interruption for the rest of the day. Example (173) in its context seems to refer primarily to duration.

At the moment, there is no settled policy about what to do with expressions like A LOT and LITTLE. We may tag some instances as FREQUENCY and others as DURATION, as appropriate, or we may introduce a new FE, say, AMOUNT, to apply to the non-FREQUENCY uses of A LOT.

A.9 Constructionally induced interpretations

A.9.1 Constructions not labeled as Extra-Thematic FEs

In some cases, there are constructions that we don't label as Extra-Thematic FEs of a target. The primary cases noted so far are speech-act (174) and epistemic (175) constructions, which combine freely with a virtually unlimited set of targets and take a variety of different FEs as their arguments.

(174) If you're still hungry, **there's** turkey in the fridge.

(175) If their teeth were long and sharp, they **ate** meat.

[3] In cases of quantifier float such as (166), the quantified phrase bears the same label as the constituent that it quantifies over. Since in (166), the object NP *them* bears the FE label THEME relative to the verb *acquire* in the Getting frame, the quantifier *all* also bears the FE label THEME.

[4] The clause *I **sharpened** it [a lot]* may refer to the degree of sharpness. In context, however, example (169) refers to the frequency of sharpening.

In each of these examples, the clause headed by *if* expresses a *CONDITION of a covert predication, in (174) a speech act, in (175) a chain of reasoning. In these cases, we now take it to be improper to annotate the *CONDITION in such examples with respect to a target in the main clause, since the frame element really belongs to a predicate that contains the main clause as a separate frame element.

A.9.2 Speech Act Construction

This is a construction which evokes the Statement frame. It takes the main clause (whether a statement or question) as an utterance and labels it as MESSAGE, binds the actual speaker of the sentence to the role SPEAKER, and then allows the expression of a number of other frame elements, including EXPLANATION (formerly REASON) (177), *CONDITION, *CONCESSIVE (178), and PURPOSE (179). This construction is used, in effect, to express the SPEAKER's reason for making a particular speech act at a particular point in the exchange, as in (176).

(176) There's some fried rice **in** the kitchen, [since you're hungry].

Mostly, such phrases have been left unannotated, but in the case of the former FE REASON especially, we have incorrectly labeled them.

(177) I promised not to reveal this, but [since you're holding a gun to my head], Kathy **took** your stapler.

(178) [Though you might not be interested], a swap-meet's gonna **happen** here next week.

(179) [Just so you know], 666 is **embedded** in every UPC barcode.

In the above examples, the indicated phrases are *not* properly frame elements of the target and, hence, should not have been labeled relative to it.

A.9.3 Epistemic Construction

This construction evokes the Evidence frame, with the main clause as the PROPOSITION and a subordinate EXPLANATION, CIRCUMSTANCES, or *CONDITION FE construed (on the second layer) as a SUPPORT which allows or demands the truth of the PROPOSITION. Consider examples (180) and (181). Example (181) belongs to the Likelihood frame, where the only REASON frame elements that occur are of this (incorrectly annotated) evidential type.

(180) John hasn't **left** yet [because his car is in the garage].

(181) It is **probable** that he was indeed in a state of shock, [since the nature of shock is to suspend ordinary or familiar reactions].

A.9.4 Second-layer annotation of certain contextually induced interpretations

One test by which we can identify phrases as CIRCUMSTANCES is to try to compound the nominal head of the candidate phrase into a phrase headed by the noun *condition*. Such paraphrases are not possible with all phrases of the same ontological type, as the contrast between (182) and (183) shows.

(182) This plant **grows** well [in sandy soil] => This plant **grows** well [in sandy soil conditions].

(183) I **met** your mother [in France] =/=> I **met** your mother [in France/French conditions].

Another test consists of presenting the candidate semi-causal state of affairs first and then resuming it with the phrase UNDER THOSE/THESE CONDITIONS (or something similar), as seen below.

(184) [At low temperatures], the metal **becomes** very fragile. [Under these conditions] it also conducts poorly.

(185) I met your mother in France. #[Under these conditions], we fell in **love**.

(186) The animals have been cared for by humans as infants. [Under such conditions], orangutan reintroduction **takes** a long time.

Another test that works for a subset of CIRCUMSTANCES is to try to form a conditional clause with the TIME, PLACE, etc., phrase part of the protasis clause and the rest of the original clause forming the main clause, as shown in the pair of examples (187) and (188).

(187) Most begonias **grow** well [in partial shade].

(188) If they are in in partial shade, most begonias grow well.

While these tests allowed us to reliably detect special semi-causal, concessive, or conditional interpretations of TIME, PLACE, etc., phrases, study of the annotated examples and reflection on the system of peripheral and extra-thematic FEs led us to conclude that these interpretations arise pragmatically (i.e. there is no overt lexical/constructional material that reliably signals their use), and that the basic use of the frame elements should not require such implications. When such implications are present, we will record them via second layer annotation with one of the extra-thematic frame elements EXPLANATION, *CONCESSIVE or *Condition. Thus, in example (182), the phrase *in sandy soil* would carry a peripheral PLACE label on the first annotation layer and the extra-thematic *CONDITION label on the second FE annotation layer. Similarly, in (189), the inspiration for (186), the *when*-clause would be labeled as the peripheral FE TIME on the first layer and with a *CONDITION label on the second annotation layer.

(189) But Kuenzer points out that even orangutan reintroduction **takes** a long time [when the animals have been cared for by humans as infants].

Let us now elaborate on the two major contexts that induce a semi-causal interpretation of TIME, PLACE, and CIRCUMSTANCES phrases. Contrast is one kind of context that promotes a causal reading of TIME, PLACE, and CIRCUMSTANCES phrases, as shown in (190) and (191).

(190) We **traveled** {slow} {in the nebula}. **Travel** was {quicker} {between Arcturus and Betelgeuse}.

(191) {During the Depression}, he **got** a {lot} of money for them. {Afterwards}, he **got** very {little}, and a guy claimed he had to pay to sell 'em.

Although there are many different uses for contrast (e.g. *I gave Harry the book and Sally the CD*), in some cases, especially when one of the contrasted elements denotes a scalar and the other denotes something easily interpretable as independently manipulable, contrast can give rise to a facilitative interpretation, which is the case in (190) and (191), above. Of course, the contrast set need not be explicitly evoked, as shown in (192).

(192) A: How much do you think are all these carpets worth?–B: Not sure. But I know that my uncle **paid** very {little} money for them {during the Depression}.

The second major context in which TIME, PLACE, and PARAMETER phrases are often interpreted as specifying EXPLANATIONs are generic sentences, such as (189), above, which describes how the world works. Although in some cases generics are definitional, generally generic sentences often imply a facilitatory or causative effect of implicit or explicit PLACE, TIME, or CIRCUMSTANCES etc., on the occurrence of the event, or more specifically the occurrence of the event with a particular subset of the participants. However, generic sentences do not automatically imply that a TIME, PLACE, or CIRCUMSTANCES phrase has an additional causal, concessive, or conditional force. To illustrate, (193) simply reports regularity, but it does not imply that the conditions facilitated or caused the regularity of behavior.

(193) [In Singapore Place], they **drive** on the left. =/=> [Under Singaporean conditions], they **drive** on the left.

Bibliography

Burchardt, A., Erk, K., Frank, A., Kowalski, A., Padó, S., and Pinkal, M. (2009). Using FrameNet for the semantic analysis of German: Annotation, representation, and automation. In Boas, H. C., editor, *Multilingual FrameNets in Computational Lexicography*, pages 209–244. Mouton.

Erk, K. and Padó, S. (2006). Shalmaneser – a flexible toolbox for semantic role assignment. In *Proceedings of the fifth International Conference on Language Resources and Evaluation (LREC-2006)*, Genoa, Italy.

Fillmore, C. J. (1968). The case for case. In Bach, E. and Harms, R., editors, *Universals in Linguistic Theory*. Holt, Rinehart & Winston, New York.

Fillmore, C. J. (1977). The case for case reopened. In Cole, P. and Sadock, J., editors, *Syntax and Semantics: Grammatical Relations*, volume 8, pages 59–81. Academic Press, New York.

Giuglea, A.-M. and Moschitti, A. (2004). Knowledge discovery using FrameNet, VerbNet and PropBank. In *Proceedings of the Workshop on Ontology and Knowledge Discovering*, Pisa, Italy. ECML.

Lakoff, G. and Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press, Chicago.

Litkowski, K. (2004). Senseval-3 task: Automatic labeling of semantic roles. In Mihalcea, R. and Edmonds, P., editors, *Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text*, pages 9–12, Barcelona, Spain. Association for Computational Linguistics.

Mel'čuk, I. (1996). *Lexical Functions in Lexicography and Natural Language Processing*, chapter Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon, pages 37–102. Benjamins, Amsterdam/Philadelphia.

Mohit, B. and Narayanan, S. (2003). Semantic extraction with wide-coverage lexical resources. In Hearst, M. and Ostendorf, M., editors, *HLT-NAACL 2003: Short Papers*, pages 64–66, Edmonton, Alberta, Canada. Association for Computational Linguistics.

Narayanan, S. (1999). Moving right along: A computational model of metaphoric reasoning about events. In *Proceedings of the /National Conference on Artificial Intelligence (AAAI '99)*, pages 121–128, Orlando, Florida. AAAI Press. http://www.icsi.berkeley.edu/~snarayan/met.ps.

Narayanan, S., Fillmore, C. J., Baker, C. F., and Petruck, M. R. (2002). FrameNet meets the semantic web: A DAML+OIL frame representation. In *Proceedings of the 18th National Conference on Artificial Intelligence*, Edmonton, Alberta. AAAI.

Pustejovsky, J. (1995). *The Generative Lexicon*. The MIT Press.

Scheffczyk, J., Baker, C. F., and Narayanan, S. (2006). Ontology-based reasoning about lexical resources. In Oltramari, A., editor, *Proceedings of ONTOLEX 2006*, pages 1–8, Genoa. LREC.

Scheffczyk, J., Borghoff, U. M., Rödig, P., and Schmitz, L. (2003). Consistent document engineering: formalizing type-safe consistency rules for heterogeneous repositories. In *DocEng '03: Proceedings of the 2003 ACM symposium on Document engineering*, pages 140–149, New York, NY, USA. ACM Press.

Scheffczyk, J., Borghoff, U. M., Rdig, P., and Schmitz, L. (2004). Towards efficient consistency management for informal applications. *International Journal of Computer & Information Science (IJCIS)*, 5(2):109–121.

Scheffczyk, J. and Ellsworth, M. (2006). Improving the quality of framenet. In Krauwer, S. and Quasthoff, U., editors, *Proceedings of the Workshop on Quality assurance and quality measurement for language and speech resources*, pages 8–13, Genoa, Italy. LREC.

Shi, L. and Mihalcea, R. (2005). Putting the pieces together: Combining FrameNet, VerbNet, and WordNet for robust semantic parsing. In *Proceedings of the Sixth International Conference on Intelligent Text Processing and Computational Linguistics*, Mexico.

Index